Circuit Simulation Including Full-Wave Maxwell's Equations

Modeling Aspects and Numerical Analysis

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)
im Fach Mathematik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
M. Sc. Christian Strohm

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät:
Prof. Dr. Elmar Kulke

Gutachter/innen:
1. Prof. Dr. Caren Tischendorf
2. Prof. Dr. Sebastian Schöps
3. Prof. Dr. Wilhelmus H. A. Schilders

Tag der mündlichen Prüfung: 22.12.2020

Zusammenfassung

Sowohl in der Elektrotechnik als auch in der Elektronik ist es üblich, elektromagnetische Bauelemente mit Hilfe von idealisierten Schaltungselementen zu modellieren. Mit der industriellen Nachfrage nach kleineren Geräten und gleichzeitig höheren Betriebsfrequenzen verlieren diese Ersatzschaltkreise ihre physikalische Rechtfertigung und Genauigkeit. Abhilfe kann geschaffen werden, indem diese Elemente durch sogenannte verfeinerte Modelle ersetzt werden, die z.B. auf den Maxwell-Gleichungen basieren und auch räumlich verteilte Effekte berücksichtigen. Die resultierenden gekoppelten Feld-/Schaltungssysteme werden in der Regel durch differential-algebraische Gleichungen beschrieben, die in dieser Arbeit unter Berücksichtigung von Modellierungsaspekten numerisch analysiert und simuliert werden.
Wir betrachten Schaltungen, die mit Hilfe der modifizierten Knotenanalyse modelliert werden, und führen ein Mock-Element ein, um verfeinerte Modelle, wie z. B. für elektromagnetische Bauelemente, Halbleiter usw., als stromsteuernde Schaltungselemente interpretieren zu können. Das verfeinerte Modell des elektromagnetischen Bauelements wird durch den vollständigen Satz der Maxwell-Gleichungen in der Lorenz-geeichten $\mathbf{A} - \boldsymbol{\varphi}$ Formulierung beschrieben, die mit Hilfe der weit verbreiteten Finite-Integrations-Technik räumlich diskretisiert werden. Es werden verschiedene Varianten für die Kopplung der elektromagnetischer Bauelemente unter Verwendung dieser Mock-Element-Schnittstelle eingeführt. Eine numerische Analyse erweitert die topologischen Kriterien für den Index der resultierenden differential-algebraischen Gleichungen, wie sie bereits in anderen Arbeiten mit ähnlichen Feld/Schaltkreis-Kopplungen hergeleitet wurden.
Um verfeinerte Modelle aus kommerzieller Software zu unterstützen, stellen wir ein Framework zur Verfügung, welches von Schaltungssimulatoren, basierend auf modifizierter Knotenanalyse, genutzt werden kann, um die Simulation dieser gekoppelten Systeme zu ermöglichen. In diesem Zusammenhang diskutieren wir Zeitintegration, Skalierungsmethoden, strukturelle Eigenschaften und einen hybriden Ansatz zur Lösung der zugrundeliegenden linearen Gleichungssysteme, der den Einsatz spezialisierter Löser für die jeweiligen Teilsysteme erlaubt.
Für die Simulation untersuchen wir sowohl einen monolithischen Ansatz als auch Waveform-Relaxationsmethoden. Im Gegensatz zum monolithischen Ansatz unterscheiden sich dabei verschiedene gekoppelte Systemformulierungen. Da für die hier betrachteten vollen Maxwell-Gleichungen die Kopplungsstruktur zusätzliche Ableitungen beinhaltet, sind bisher existierende Konvergenzaussagen für die Waveform-Relaxation von gekoppelten differential-algebraischen Gleichungen nicht anwendbar. Wir präsentieren hier eine erweiterte Konvergenzanalyse mit hinreichenden Kriterien für die Konvergenz der Waveform-Relaxation von differential-algebraischen Gleichungssystemen, deren Kopplungsterme auch Ableitungen enthalten. Darauf aufbauend entwickeln wir hinreichende topologische Kriterien, die eine Konvergenz von Gauß-Seidel- und Jacobi-artigen Waveform-Relaxationen für die gekoppelten Feld/Schaltkreis-Modelle garantieren.
Darüber hinaus stellen wir numerische Benchmarks zur Verfügung, die die eingeführten Me-

thoden und Theoreme dieser Abhandlung unterstützen. Diese Benchmarks werden mit Hilfe eines selbstimplementierten Schaltungssimulators erstellt, der sowohl eine selbstimplementierte Version eines Modellierers für elektromagnetische Bauelemente als auch ein industrielles Werkzeug miteinander verbindet. Der Erfolg der monolithischen Simulation wird für Testfälle mit elektromagnetischen Bauelementen und Halbleitern verifiziert. Weitere Benchmarks zeigen die Konvergenz der Waveform-Relaxationsmethoden, wo sie zu erwarten ist, und Divergenz, wo hinreichende Kriterien verletzt werden.

Abstract

In electrical or electronic engineering, it is common to model electromagnetic devices by means of lumped circuit elements. With the industrial demand for smaller devices and simultaneously higher operating frequencies, current parameter models lose their physical justification and accuracy. This can be remedied by replacing the critical elements in the surrounding circuits with refined models based, for example, on Maxwell's equations, which take spatially distributed effects into account. The resulting coupled field/circuit systems are usually described by differential-algebraic equations, which are numerically analyzed and simulated in this thesis under consideration of modeling aspects.
We consider circuits which are modeled using modified nodal analysis and introduce a mock element to interface refined models such as electromagnetic devices, semiconductors etc., by interpreting them as current controlling lumped elements. The refined model of the electromagnetic field device is described by the full set of Maxwell's equations given in Lorenz-gauged $\mathbf{A} - \boldsymbol{\varphi}$ formulation, which are spatially discretized using the widely spread finite integration technique. Different variants are introduced for the coupling of electromagnetic field devices into the circuit using this mock element interface. A numerical analysis extends the topological criteria for the index of the resulting differential-algebraic equations, as already derived in other works with similar field/circuit couplings.
In order to support discretized refined models which can also be black boxes, we provide a framework that can be used by circuit simulators, based on modified nodal analysis, to enable the simulation of these coupled systems. In this context we discuss time integration, scaling methods, structural properties and a hybrid approach to solve the underlying linear systems of equations, which allows the use of specialized solvers for the respective subsystems.
For the simulation we study both, a monolithic approach and waveform relaxation methods. In contrast to the monolithic approach, various coupled system formulations differ when waveform relaxation methods are applied to them. Since the coupling structure contains additional derivatives for the full Maxwell equations considered here, previously existing convergence statements for the waveform relaxation of coupled differential-algebraic equations are not applicable. We present here an extended convergence analysis with sufficient criteria for the convergence of the waveform relaxation of differential-algebraic equation systems whose coupling terms also contain derivatives. Based on this, we develop sufficient topological criteria which guarantee convergence of Gauss-Seidel and Jacobi type waveform relaxation schemes for the coupled field/circuit models.
Furthermore, we provide numerical benchmarks which support the methods and theorems introduced in this treatise. These benchmarks are produced using a self-implemented circuit simulator interfacing both, a self-implemented version of an electromagnetic device modeler and an industrial tool. The monolithic simulation's success is verified for test cases with electromagnetic field devices and semiconductors. Further benchmarks demonstrate convergence of

waveform relaxation methods where it is to be expected and divergence where sufficient criteria are violated.

Preface

First of all I would like to thank my supervisor Prof. Dr. Caren Tischendorf, who made it possible for me to work on my thesis, always supported me and with whom I had many fruitful discussions. Since the beginning of my bachelor thesis, she has always accompanied me and encouraged me whenever there was a challenge. Furthermore, I would like to thank Prof. Dr. Sebastian Schöps, not only for his willingness to serve as a referee, but also for his helpfulness and the pleasant cooperation. In the course of my work on the project nanoCOPS I had the chance to meet Prof. Dr. Wilhelmus Schilders, to whom I am indebted that he agreed to be a referee.

This thesis was written during my time in the science group and chair of Prof. Dr. Caren Tischendorf where I worked on the EU founded Seventh Framework Programme ICT project nanoCOPS (GA619166) and the Federal Ministry for Economic Affairs and Energy founded project MathEnergy (0324019E) to which I am grateful. I had the opportunity to meet many extraordinary people who inspired me not only in a scientific but also in a personal way, some of them became my friends. I especially appreciate the fact that Dr. Lennart Jansen, Dr. Sascha Baumanns and Dr. Christoph Huck, for example, were always available even after they had been my colleagues. In reality, my gratitude is due to many people, not only to my colleagues in Berlin but also to Dr. Idoia Cortes Garcia, Dr. Wim Schoenmaker, Dr. Peter Meuris, Prof. Dr. Hans Georg Brachtendorf and Dr. Kai Bittner for scientific discussions and joint publications.

Besides the support of my colleagues, I have also profited a lot from my friends. Manuel Radons was always available with advice and action. With Tom Streubel I often had hours of scientific sessions. Of course, I also want to thank my great family, especially my outstanding mother, who made it possible for me to move from Cologne to Berlin, where I started my doctoral studies.

But without a doubt there is one person in this world who definitely enjoys my greatest gratitude. Lovable, caring, insightful, wise and strong, I am talking about my exceptionally captivating wife Dasha who has been an anchor and muse to me since our first encounter.

Я тебя люблю моя буся

Christian Strohm

Berlin, 14. August 2020.

Contents

Notation xi

List of Figures xv

List of Tables xvii

1 Introduction 1

2 Differential-Algebraic Equations 5
- 2.1 Preliminaries . . . 6
- 2.2 Coupled Systems . . . 9
- 2.3 The Dissection Index . . . 10
- 2.4 Conclusions . . . 13

3 Electric Circuits 15
- 3.1 Matrix Representation and Characterization of Circuit Structures . . . 16
- 3.2 Lumped Circuit Modeling . . . 20
- 3.3 Decoupling of the Lumped Circuit DAE . . . 26
- 3.4 Conclusions . . . 33

4 Electromagnetic Devices 35
- 4.1 Maxwell's Equations . . . 37
- 4.2 Constitutive Relations . . . 40
- 4.3 Interface and Boundary Conditions . . . 43
  - 4.3.1 Media Interface Conditions . . . 43
  - 4.3.2 Spatial Boundary Conditions . . . 46
- 4.4 $\mathbf{A} - \boldsymbol{\varphi}$ Formulation . . . 49
  - 4.4.1 Boundary Conditions . . . 51
  - 4.4.2 Gauge Fixing . . . 51
  - 4.4.3 Two System Formulations . . . 52
- 4.5 Spatial Discretization using Finite Integration Technique . . . 54
  - 4.5.1 Discretization through Cell Complex . . . 55
  - 4.5.2 Duality . . . 58
  - 4.5.3 Orientation . . . 61
  - 4.5.4 Maxwell's Grid Equations . . . 64
  - 4.5.5 Constitutive Grid Relations . . . 71
  - 4.5.6 Boundary Grid Conditions . . . 74
  - 4.5.7 Maxwell's Grid Equations in Potential Formulation . . . 75

4.6 Conclusions . . . 79

5 Coupled Electric Circuits and Electromagnetic Field Devices 81
5.1 Coupled Modeling . . . 82
5.1.1 Coupling Branches . . . 83
5.1.2 Coupling Potentials . . . 84
5.1.3 Coupling Currents . . . 85
5.2 Coupled System Formulations . . . 87
5.2.1 EM Device as Mock Element . . . 87
5.2.2 Resulting Coupled Systems . . . 93
5.3 Analysis of Coupled Field/Circuit DAEs . . . 97

6 Simulation of Field/Circuit Coupled Systems 109
6.1 Interface Realization . . . 111
6.2 Monolithically Solving the Field/Circuit Coupled Problem . . . 112
6.2.1 Time Integration . . . 113
6.2.2 Newton's Method . . . 115
6.2.3 Linear Solving Methods . . . 116
6.2.4 Scaling Issues . . . 119
6.2.5 Structural Exploitation . . . 122
6.3 Waveform Relaxation Method . . . 123
6.3.1 Gauss-Seidel Type for Field/Circuit Coupled System . . . 126
6.3.2 Jacobi Type for Field/Circuit Coupled System . . . 132
6.3.3 Convergence Analysis . . . 134
6.3.3.1 Convergence of Gauss-Seidel Type GS1 . . . 141
6.3.3.2 Convergence of Gauss-Seidel Type GS2 . . . 148
6.3.3.3 Convergence of Jacobi Type Jac1 . . . 150
6.3.3.4 Convergence of Jacobi Type Jac2 . . . 151
6.4 Conclusions and Remarks . . . 151

7 Numerical Benchmarks 153
7.1 Monolithic Simulation Benchmarks . . . 153
7.1.1 Coupled Field/Circuit Systems . . . 153
7.1.2 Using `DevEM` Models . . . 156
7.2 Waveform Relaxation Benchmarks . . . 162
7.2.1 Low-Pass Filter with Current Source . . . 162
7.2.2 Band-Pass Filter using Capacitor and Inductor . . . 167
7.3 Conclusions . . . 173

8 Summary and Outlook 175

A Appendix 181
A.1 Electromagnetism . . . 181
A.2 Graph Theory . . . 181
A.3 Functional Analysis . . . 184

A.4 Convergence Analysis 187

Bibliography **197**

Index **217**

Notation

Abbreviations

MNA	modified nodal analysis
ODE	ordinary differential equation
DAE	differential-algebraic equation
WR	waveform relaxation
PDE	partial differential equation
EM	electromagnetic
MQS	magneto-quasistatic
EQS	electro-quasistatic
FIT	finite integration technique
FDTD	finite difference time domain
MEs	Maxwell's equations
GL	Gauss's law
GLM	Gauss's law for magnetism
MF	Maxwell-Faraday's law
MA	Maxwell-Ampère's law
CCE	current continuity equation
PEC	perfectly electrically conducting
MGEs	Maxwell's grid equations
CGRs	constitutive grid equations
BDF	backward differentiation formula
RK	Runge-Kutta
C	capacitor
R	resistor
L	inductor
V	voltage source
I	current source
M	mock element
E	electromagnetic field device
X^+	contains at least one element of type X
GS1	first Gauss-Seidel type waveform relaxation class
GS2	second Gauss-Seidel type waveform relaxation class
Jac1	first Jacobi type waveform relaxation class
Jac2	second Jacobi type waveform relaxation class

General

$\exists$	it exists
$\exists!$	it exists exactly one
$\forall$	for all
$\emptyset$	empty set
$\mathbb{K}$	any set of numbers
$\mathbb{N}$	integers
$\mathbb{R}$	real numbers
$\mathbb{R}^m$	m-dimensional vector space of real numbers
$\mathbb{R}^{\mathcal{N}}$	family of real numbers indexed by elements in $\mathcal{N}$
$C^p(\mathcal{A},\mathcal{B}), C^p(\mathcal{A}), C^p$	linear space of p-times continuously differentiable functions mapping from $\mathcal{A}$ to $\mathcal{B}$, open
$A \in \mathcal{A}$	A is an element (vector, function, matrix etc.) of set $\mathcal{A}$
$A\|_{\mathcal{B}}$	A restricted to domain $\mathcal{B}$
$\mathcal{A} \subset \mathcal{B}$	set $\mathcal{A}$ is contained in set $\mathcal{B}$
$\mathcal{A} \not\subset \mathcal{B}$	set $\mathcal{A}$ is not contained in set $\mathcal{B}$
$\mathcal{A} \cap \mathcal{B}$	intersection of sets $\mathcal{A}$ and $\mathcal{B}$
$\mathcal{A} \cup \mathcal{B}$	union of sets $\mathcal{A}$ and $\mathcal{B}$
$\mathcal{A} \backslash \mathcal{B}$	set of elements in $\mathcal{A}$ but not in $\mathcal{B}$
$\mathcal{A} \times \mathcal{B}$	product set of sets $\mathcal{A}$ and $\mathcal{B}$
$\mathcal{A} \to \mathcal{B}$	maps from set $\mathcal{A}$ to set $\mathcal{B}$
$\mathcal{A} \oplus \mathcal{B}$	direct sum of sets $\mathcal{A}$ and $\mathcal{B}$
$\|\mathcal{A}\|$	number of elements in set $\mathcal{A}$
$\{A, B\}$	set consisting of elements A and B
$\ker A$	kernel of A
$\operatorname{im} A$	image of A
$\operatorname{rank} A$	rank of A
$\|A\|$	magnitude or absolute value of A
$\|\|A\|\|$	norm of A
A^{-1}	inverse of A
$A^\top$	transposed of A
$\rho(A)$	spectral radius of matrix A
I	identity matrix
$\mathcal{I} = [t_0, T]$	time interval starting from t_0 and ending with T
$\mathcal{D}$	domain of definition
Ω	spatial domain of consideration
$\partial\Omega$	boundary of spatial domain Ω
$\bar{\Omega}$	closure of domain Ω
Γ	boundary surface or interface
n	unit normalvector if not declared otherwise
τ	unit tangential vector if not declared otherwise
$x \cdot y$	dot product of vectors x and y
$x \times y$	cross product of vectors x and y
x'	jet-derivative operator acting on x

$\frac{\partial}{\partial t}x$	partial derivative of x with respect to t
$\frac{\mathrm{d}}{\mathrm{d}t}x$	total derivative of x with respect to t
$\int_{\mathcal{X}} x(t)\,\mathrm{d}t, \int_a^b x(t)\,\mathrm{d}t$	integral of function x over the domain $\mathcal{X}$ or from a to b
$\int, \iint, \iiint$	integrals over domains of respectively dimensions 1,2 and 3
$\oint, \oiint$	closed domains integrals of respectively dimensions 1 and 2
$\int_{\mathcal{X}} x(t) \cdot \mathrm{d}t$	integral's body first performs a dot product with point-wise unit normal/tangential vector
$\max_{t\in\mathcal{X}}\{x(t)\}$	maximum of comparable $x(t)$ over all t in set $\mathcal{X}$
$\sup_{t\in\mathcal{X}}\{x(t)\}$	supremum of comparable $x(t)$ over all t in set $\mathcal{X}$
$\sum_{t\in\mathcal{X}} \mathrm{x(t)}$	summation of $x(t)$ over all elements t in countable set $\mathcal{X}$
$\bigcup_{t\in\mathcal{X}} x(t)$	Union of sets $x(t)$ over all elements t in countable set $\mathcal{X}$
$\lim_{k\to\infty} x^{[k]}$	the limit value of $x^{[k]}$ for k against infinity
$\lambda(\mathcal{X})$	Lebesque measure of Lebesque-measurable $\mathcal{X}$
∇	gradient operator
$\nabla\cdot$	divergence operator
$\nabla\times$	curl operator
∇^2	vector Laplace operator
Δ	infinitesimal small

List of Figures

1.1 Low-pass filter . . . 2

2.1 η-DAE motivating proper stating derivative term . . . 8

3.1 Multi-terminal circuit element . . . 16
3.2 Kirchhoff's circuit laws . . . 21
3.3 MOSFET equivalent circuit . . . 23
3.4 Electric circuit example . . . 26

4.1 EM device example - balun . . . 37
4.4 Interface condition visualization . . . 45
4.5 Possible boundaries and interfaces . . . 45
4.6 Maxwell's house (continuous) . . . 50
4.7 Rectilinear grid . . . 57
4.8 p-cells for rectilinear grid . . . 58
4.9 Staggered dual cell complex: mesh points and volumes . . . 61
4.10 Staggered dual cell complex: facets . . . 62
4.11 Staggered dual cell complex: links . . . 63
4.12 Example of staggered meshes on Ω . . . 64
4.13 p-cells orientation for rectilinear grid . . . 65
4.14 FIT visualisation of GLM . . . 70
4.15 FIT visualisation of MF . . . 71
4.16 Constitutive grid relation 1 . . . 74
4.17 Constitutive grid relation 2 . . . 75
4.18 Maxwell's house (discrete) . . . 76

5.1 Electric circuit incorporating EM device . . . 81
5.2 Electric circuit incorporating EM device as mock element . . . 90
5.3 Electric circuit incorporating EM device partially as mock element . . . 92
5.4 voltage source coupled EM device . . . 106
5.5 voltage source coupled EM device - circuit subsystem of shifted variant . . . 107

6.1 Coupling schemes . . . 109
6.2 Quasi-canonical momentum elimination . . . 111

7.2 Simulation plot low-pass filter with EM device . . . 154
7.1 Low-pass circuit with EM device . . . 154
7.3 Quasi-canonical momentum elimination . . . 156

7.4 Low-pass circuit with EM device from DevEM . . . 156
7.6 Balun test circuit with EM device from DevEM . . . 157
7.5 Simulation plot low-pass filter with EM device from DevEM . . . 158
7.7 Simulation plot balun circuit . . . 159
7.8 Balun sectional view . . . 159
7.10 Simulation plot band-pass filter with DevEM devices . . . 160
7.9 Band-pass circuit with EM devices from DevEM . . . 160
7.11 Band-pass filter . . . 161
7.12 CMOS inverter with semiconductor . . . 161
7.13 Current low-pass filter coupled EM . . . 163
7.14 Current low-pass filter coupled EM - divergence of GS1 and Jac1 . . . 164
7.15 Current low-pass filter coupled EM (modified) . . . 164
7.16 Current low-pass filter coupled EM (modified) - convergence of GS1 and Jac1 . . . 165
7.17 Current low-pass filter coupled EM (modified) - solution near zero variable . . . 165
7.18 Current low-pass filter coupled EM (modified) - error near zero variable . . . 166
7.19 Current low-pass filter coupled EM (modified) - solution node potential . . . 166
7.20 Current low-pass filter coupled EM (modified) - error node potential . . . 167
7.21 Band-pass filter using capacitor and inductor . . . 167
7.22 Band-pass filter coupled EM variant 1 . . . 168
7.23 Band-pass filter coupled EM variant 1 - convergence of Gauss-Seidel . . . 169
7.24 Band-pass filter coupled EM variant 1 - solution node potential . . . 170
7.25 Band-pass filter coupled EM variant 1 - error node potential . . . 170
7.26 Band-pass filter coupled EM variant 1, windowing - convergence of Gauss-Seidel 171
7.27 Band-pass filter coupled EM variant 1, windowing - solution node potential . . . 171
7.28 Band-pass filter coupled EM variant 1, windowing - error node potential . . . 172
7.29 Band-pass filter coupled EM variant 2 . . . 172
7.30 Band-pass filter coupled EM variant 2 - divergence of Gauss-Seidel . . . 173
7.31 Band-pass filter coupled EM variant 2 - error node potential . . . 174

8.1 Power-to-gas coupled network . . . 177

List of Tables

3.1 SI units of electric circuit quantities 20
3.2 Basic two-terminal circuit elements 22

4.1 Maxwell's equation 38
4.2 Maxwell's equations in integral form 38
4.3 SI units of EM quantities 39
4.4 SI units of differential operators 39

6.1 Quantity scales 121

7.1 Material properties 153
7.2 Monolithic simulation of Example 7.1 with GMRES 155
7.3 Errors for monolithic simulation of Example 7.1 using `gmres` 155
7.4 Errors for monolithic simulation of Example 7.1 using `spsolve` 155
7.5 Refined model test cases for `DevEM`. 157

1 Introduction

Innovations in the field of electromagnetic devices are one of the most important driving forces in our society. With the help of computer aided design tools, pioneered by electric circuit simulation, a continuously increasing advancement could be ensured while saving expenditures by virtue of prototype manufacturing. Keeping pace with industrial demands becomes more challenging especially with regard to the development of integrated circuits. To be more precise, the aggressive scaling of device lengths, simultaneous increase of operating frequencies and possible heat influence call for more complex models. At the same time, efficient and robust numerical solution methods are needed to cope with the complexity especially when simulating in time-domain.

In both electrical and electronic engineering, circuits are modeled in terms of lumped elements which are assumed to behave in an idealized way that is neglecting spatially distributed phenomena such as cross-talking or skin effect. In order to integrate more complex elements, such as transistors, into the circuit, so-called equivalent circuits were used for a long time. They have the function to imitate the behavior of complex elements by means of classical lumped ones such as capacitors, inductors, resistors and sources. With the increasing number of parameters in the hundreds, it was soon understood that these equivalent circuits no longer had any physical justification. A circumvention of these problems is possible with the help of refined models based on Maxwell's equations since they describe all macroscopic electromagnetic phenomena.

A most simple low-pass filter shows already a big difference when dealing with Gigahertz frequencies, see Figure 1.1.

This field/circuit coupled methodology enables some possibilities, especially in the semiconductor industry where it provides a great tool to test newly designed devices, described by electromagnetic field models, in the larger circuit environment they are intended for usage. Contrary to treating the device as an input/output system, e. g. observing the electric current at some contact given a certain voltage bias, the engineer now has the possibility to validate the functionality in addition to the physical performance alone.

Within the *nanoCOPS*[1] project, industrial and academic partners aimed to advance a methodology of circuit-and-system-level modeling and simulation, especially with regard to RF-circuity in wireless communication involving field/circuit/heat coupling. This treatise is devoted to the coupled field/circuit modeling and simulation. Various coupling formulations are presented, analyzed and their consequences for the simulation process upon topological design decisions are discussed.

[1] EU funded FP7 ICT project nanoCOPS (GA619166)

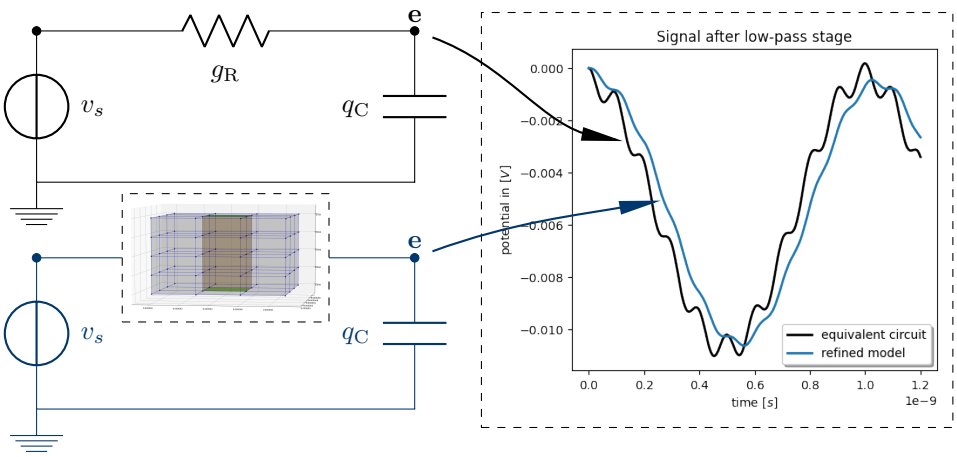


Figure 1.1: A low-pass filter with a signals of equal strength at 1 GHz and 10 GHz. With a resistor as an equivalent circuit (up) and the according refined model (down).

Circuit Modeling and Simulation The most popular modeling approach in commercial circuit simulators is the modified nodal analysis (MNA), cf. [HRB75; CL75; DK69; CDK87; FG05; RS09]. Given either in conventional or charge oriented formulation, the MNA considers capacitors, inductors, resistors, voltage sources and current sources as lumped elements. The gathered circuit model's equations forms a system of differential-algebraic equations (DAE). DAEs are an excellent modeling tool when differential and algebraic equations, or systems thereof, are involved, see for example the applications in [BCP89; UKM95; RR02; OOL04; SM05; KM06]. It is known that the simulation of problems based on DAEs is more difficult than that based on systems of ordinary differential equations (ODEs), especially the higher their index is. Due to the strong relation of circuits with graph theory, topological criteria allow for the deduction of this structural information. In general, the index does not exceed 2 and the equations, causing critical perturbation of the system, only contribute in a linear way, see e. g. [Tis99; ET00]. As the system sizes for circuits are often very large, waveform relaxation (WR) methods have been of interest for about 35 years ago. Convergence results have been proven for WR schemes solving ODEs [Lel82; LRS82; MN87; Bur95] and later for DAEs [JK96; AG01] or for distributed models in [GS98].

Field Modeling and Simulation When modeling multidimensional problems, one likely encounters partial differential equations (PDEs), for instance Maxwell's Equations (MEs) when describing physical phenomena [Max65; Jac99]. Electromagnetic (EM) field device modeling also falls within this category with quantities depending on space and time. Here, a set of PDEs is chosen so that the physical phenomena are sufficiently described. Depending on the application, one usually starts with MEs or alternative formulations thereof, often also with respect to simplifications, see e. g. [Car80; HM89; Bír+90; Cle+02; WH06; Ban10; Che+11; MNS17; Cor20] and may more. In case of semiconductors, we may encounter further PDEs describing fluid dynamic, kinetic or quantum effects in the materials, see e. g. [Sel84; RSM90; Arg92; Arn+04; DMR05].

Typically, as the next step, these PDEs are then spatially discretized, e. g. by finite difference time domain, finite element, cell method or finite integration technique (FIT), see further [Yee66; Bos98; Ton01; Wei77a], always resulting in a system described by an ODE or DAE whereby time is the remaining dimension. Not only because of the enormous size of the systems, which justifies the need for new numerical solution methods, but also because of structural exploitation possibilities, the development of tailor-made solution methods has been addressed, e. g. multi-grid [Hac13; Mer+98] Krylov subspace [Saa03] or Tikhonov regularization methods [Cal+00; Bru+17]. Another very popular approach is the model order reduction where one tries to downsize the systems before time integration is performed, see e. g. [SVR08; BS17; RS09; Alì+13].

Coupled Field/Circuit Modeling and Simulation Usually, for more complex elements in circuits it is a common approach to use equivalent circuits. As their accuracy and physical interpretation is limited, in particular with regard to current industrial requirements, the interest in refined modeling for circuitry is growing. Hereby, elements, such as EM field devices, are modeled by sets of PDEs, sufficiently describing the physical phenomena. Spatial discretization yields an ODE or DAE whose quantities are then matched with the circuit ones, which are voltages (potentials) and currents, so that these elements can be incorporated into the circuit as lumped elements e. g. [HM76; CLP98; Tsu+93; Wan96; DGL99; DHW04; SDB10; Bau12]. Here the advantage of DAEs as a modeling tool becomes clear as the DAEs from circuitry can easily be combined with the DAEs from refined models yielding yet another DAE. For an analytical investigation of DAE-PDE coupling we refer to [Tis04; LMT05].
These field/circuit models represent challenging systems to solve. Numerical analysis concerning the index can be found for instance in [Bau12; Jan15; Cor+19] or for conductor models in [Ben06; Sch11]. Whereas it is in the nature of things themselves that specialized simulation methods have been researched for the separate physical model equations, the matter is different for coupled problems. In particular, when it comes to the solution methods of the underlying linear systems of equations, there is still much potential in structural exploitation. Here, also the co-simulation ansatz has proven to be particularly efficient and not only for different field/circuit coupled models, see e. g. [SDB10; Cle+12] for some examples. Especially WR methods have the advantage that they allow for black-box solvers on each subsystem, for instance the ones that come with the subsystems' modeler. These methods were further analyzed, for example, in [AG01; Bed93; Ebe08; Alì+12; Bar+13] with regards to the convergence rate and convergence criteria which appear to depend on coupling topology and circuit parameters, as shown for certain subsystem and coupling models in [PT18].

Motivated by the collaboration with *MAGWEL*[2] and the work with their device-electromagnetic modeler `DevEM`[3], this treatise deals with the full-wave MEs given in Lorenz-gauged $\mathbf{A} - \boldsymbol{\varphi}$ formulation and spatially discretized using FIT, see [Sch+16; Bau12]. In comparison with the similar field/circuit modeling in [Bau12], a more generic coupling using a mock element interface is introduced in favor of the MNA philosophy, cf. [ST20]. This interface allows various

[2] *MAGWEL N.V.* is one of the industry partners in the *nanoCOPS* project
[3] an industrial grade software package for the co-simulation of semiconductor devices together with metal interconnect developed by *MAGWEL N.V.* [MAG16]

variants for the coupling of EM field device into the circuit. This mock element approach is similar to the Mock-up interface considered for instance in [Blo+11].

For the simulation of coupled field/circuit models, both, a monolithic approach and waveform relaxation methods, as a popular co-simulation ansatz, are studied.
For a monolithic simulation various methods are introduced and strategies are discussed. In particular, scaling methods, discussions on direct and iterative solvers, justification of hybrid methods for solving underlying linear systems of equations and structural exploitation. The coupled model equations introduced in this treatise are considered with the prospect of allowing them to use black-box models for field devices such as those from `DevEM`. The coupled field/circuit simulations of test cases are then performed by a self-implemented electric circuit simulator interfacing a self-implemented version of an EM device modeler and `DevEM` using the mock element based interface design.

This treatise is organized as follows. We start with an introduction to DAEs in Chapter 2 where we lay down the fundamental differences to ODEs, their origin when coupling different systems and introduce the dissection index concept for later analysis. Chapter 3 is devoted to the electric circuit modeling using MNA and introduces a mock element as an interface for refined models, especially the EM field device in Chapter 4. Graph theoretical topologies and their consequences for the MNA structure are discussed. Based on this and the dissection index concept, various decoupled formulations are provided as preparation for the convergence analysis of WR methods in Chapter 6. In Chapter 4, we introduce an EM field device model as a refined model based upon the full set of MEs, which is spatially discretized using FIT. The field/circuit coupling is studied in Chapter 5. Here we derive different coupling forms using the mock element modeling interface. In Chapter 6, we deal with the numerical methods solving these field/circuit coupled systems. We consider both, a monolithic approach in Section 6.2, and WR methods as popular variant of co-simulation in Section 6.3. We discuss time integration, scaling techniques, hybrid solving strategies of high dimensional problems linear systems and structural exploitation. Furthermore, we derive a convergence theorem for waveform relaxation methods. For Gauss-Seidel and Jacobi type WR schemes, applied to the coupled field/circuit systems, we provide sufficient topological criteria guaranteeing convergence of these. In Chapter 7, we support these simulation strategies by numerical benchmarks of coupled field/circuit test cases using the self-implemented circuit simulator and EM field device modeler as well as `DevEM` interfaced models. The first part gives justification for the different methods developed for the monolithic approach and shows some simulation results with industrial prototype devices. In the second part, we perform some benchmarks on the WR criteria showing convergence where it is to be expected and divergence where criteria are violated. We conclude with Chapter 8 where we summarize results of the treatise and give an outlook.

2 Differential-Algebraic Equations

Its been a couple of decades since differential-algebraic equations (DAEs) got more and more popular mainly because of their use as a modeling tool. They are encountered in various fields of application among which are for instance the simulation of circuits, multibody mechanics, chemical reactions, cardiovascular systems, optimal control problems and many more, see e. g. [BCP89; UKM95; RR02; OOL04; SM05; KM06]. Besides this, DAEs also arise for other reasons such as for instance after semi-discretization of partial differential equation systems, which will be the case in Section 4.5.

DAEs can be understood as a class expansion of systems of ordinary differential equations (ODEs). Although DAEs differ from classical ODEs in many ways, they constitute somehow uniformly singular ODEs, see further [GM86; BCP89; LMT13]. To be more precise, consider the equation

$$f(x'(t), x(t), t) = 0 \tag{2.1}$$

defining relations for some vector valued function. If, for all (x^1, x, t) of the domain of interest, the partial Jacobian $\frac{\partial}{\partial x^1} f(x^1, x, t)$ is nonsingular, we call 2.1 a regular implicit ODE. On the other hand, if the Jacobian is singular on the domain of interest, (2.1) is said to be a DAE.

The attractiveness of DAEs now becomes clear: with this tool, model equations can be generated automatically whereby the preservation of certain mathematical structures is not required anymore. Beyond that, couplings of different models or systems are hardly an obstacle anymore.
This makes the first stage of the solving process, the modeling part, much easier, especially when the systems are of large dimensions. In exchange for the simple modeling, the difficulty lies in the second stage which is the numerical solving of the DAEs. In particular, for a DAE some components of the solution are determined by the so-called *hidden constraints*, restricting the choice of initial values. Further, DAEs may demand higher smoothness of the solution. As a consequence, numerical methods can fail whereas they are successful for ODEs, see e. g. [GM86; HW91; LMT13]. One approach to reduce the required smoothness is given by the introduction of properly stated derivative terms, see [LMT13], which we concentrate on herein.

In order to categorize a DAE in conjunction with its structural complexity and, therefore, numerical difficulty, several so-called index-concepts have been developed. There are for instance the concepts of *differentiation index* [BCP89; CG95], *perturbation index* [HLR89], *strangeness index* [KM94; KM06], *tractability index* [GM86; Mär87; LMT13] or *dissection index* [Jan15]. In any case, these concepts constructively measure the DAE's deviation from being a regular ODE and thus providing a way of decoupling the DAE, i. e. how to separate the inherent ODE, the pure/constraint free differential relation, from its algebraic dependencies. The decoupling

process is an analytical tool and not often encountered during practical solving. Instead, one often makes use of index reduction methods, for example the *Gear-Gupta-Leimkuhler stabilization* [GLG85]. In this thesis, however, we focus on the dissection index concept as it combines the strengths of both strangeness index and tractability index concepts and provides a comprehensible topological interpretation, see e. g. [PT18].

The chapter is structured as follows. First, we introduce DAEs with nonlinear derivative term, a special form of DAEs that is dealt with later, and provide some basic properties and classifications thereof. Secondly, we introduce the term coupled system as one key origin of DAEs, for instance when combining model equations of various problems such as circuit models in Chapter 3 and electromagnetic field models in Chapter 4. Finally, we briefly introduce the dissection index concept with an alternative matrix chain ending as preparation for the analysis in Chapter 5.

2.1 Preliminaries

Contrary to the just introduced DAEs in standard form (2.1), we focus on a slightly more specific form that is DAEs with derivative terms following [LMT13]. This form of DAEs comes mainly from circuit simulation where it is widely used and analyzed, see e. g. [Mär03].

Let $\mathcal{I} \subset \mathbb{R}$ and $\mathcal{D} \subset \mathbb{R}^m$ be open subsets with $t_0 \in \mathcal{I}$. Consider a continuous function $f : \mathbb{R}^n \times \mathcal{D} \times \mathcal{I} \to \mathbb{R}^m$ with continuous partial derivatives $\frac{\partial}{\partial y} f(y, x, t)$ and $\frac{\partial}{\partial x} f(y, x, t)$ and a continuous differentiable function $d : \mathcal{D} \times \mathcal{I} \to \mathbb{R}^n$.

Definition 2.1 (DAE with nonlinear derivative term)
Given that for all $(y, x, t) \in \mathbb{R}^n \times \mathcal{D} \times \mathcal{I}$ the partial derivative $\frac{\partial}{\partial y} f(y, x, t)$ is singular, we call

$$f(\frac{\mathrm{d}}{\mathrm{d}t} d(x(t), t), x(t), t) = 0 \tag{2.2}$$

a DAE with nonlinear derivative term.
A continuous function x_* satisfying pointwise (2.2) on an interval $\mathcal{I}_* \subset \mathcal{I}$, with values $x_*(t) \in \mathcal{D}, t \in \mathcal{I}_*$ such that $d(x_*(\cdot), \cdot)$ is continuously differentiable, is called a solution thereof.
The value $x_0 \in \mathcal{D}$ is said to be a consistent initial value if the initial value problem, that is (2.2) with initial condition

$$x(t_0) = x_0 \tag{2.3}$$

possesses a solution.

Remark 2.2
Note that all DAEs in standard form (2.1) can be translated into one with nonlinear derivative term (2.2) for instance by simply defining d as the identity.

Properties Contrary to ODEs, for DAEs typically not all components of the solution have to be differentiable whereas the so-called right-hand side may require higher smoothness. Further, an initial condition cannot be chosen completely arbitrary due to the algebraic constraints, see for instance the following example.

Example 2.3
For a given right-hand side function $q_2 : \mathbb{R} \to \mathbb{R}$, consider the DAE

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} x_1(t) + x_1(t) - x_2(t) &= 0, \\ x_1(t) &= q_2(t) \end{aligned}$$

whose solution is given by

$$x_1(t) = q_2(t) \quad \text{and} \quad x_2(t) = \frac{\mathrm{d}}{\mathrm{d}t} q_2(t) + q_2(t).$$

We make the following observations:

(i) Only x_1 has to be differentiable.

(ii) For an initial condition $x_0 = x(t_0)$, the initial value is constrained to

$$x_0 = (q_2(t_0), \frac{\mathrm{d}}{\mathrm{d}t} q_2(t_0) + q_2(t_0)).$$

(iii) The right-hand side is required to be at least one time differentiable.

Motivated by the first observation in Example 2.3, we introduce the concept of properly stated derivative terms.

Definition 2.4 (properly stated derivative term, [LMT13])
The DAE (2.2) has on $\mathcal{D}\times\mathcal{I}$ a properly stated derivative term, if $\ker \frac{\partial}{\partial y} f(y, x, t)$ *and* $\operatorname{im} \frac{\partial}{\partial x} d(x, t)$ *are C^1-subspaces in $\mathbb{R}^n$, and the transversality condition*

$$\ker \frac{\partial}{\partial y} f(y, x, t) \oplus \operatorname{im} \frac{\partial}{\partial x} d(x, t) = \mathbb{R}^n, \quad \forall (y, x, t) \in \mathbb{R}^n \times \mathcal{D} \times \mathcal{I},$$

holds.

Describing a problem by a DAE with properly stated leading term instead of a DAE in standard form has quite some remarkable advantages. First, the solution does not need to meet unnecessarily smoothness demands as the derivative part is precisely given, see [LMT13]. Secondly, and as a consequence, numerical methods perform better, see for instance Example 2.5.

Example 2.5
Let $\mathcal{I} \subset \mathbb{R}$ and $\eta \in \mathbb{R}$, consider the well-know DAE from [GP83] in standard form

$$\frac{\mathrm{d}}{\mathrm{d}t} x_1(t) + \eta t \frac{\mathrm{d}}{\mathrm{d}t} x_2(t) + (1 + \eta) x_2(t) = 0, \tag{2.4a}$$

$$x_1(t) + \eta t x_2(t) = \exp(-t). \tag{2.4b}$$

For values $\eta < 0.5$ the so-called η-DAE (2.4) is known to be numerically instable in that classical ODE methods, such as the implicit Euler method fail, see Figure 2.1. Rewriting (2.4) to an equivalent DAE with proper stated leading term, that is

$$\frac{\mathrm{d}}{\mathrm{d}t}x_1(t) + \frac{\mathrm{d}}{\mathrm{d}t}(\eta t x_2(t)) + x_2(t) = 0, \tag{2.5a}$$

$$x_1(t) + \eta t x_2(t) = \exp(-t), \tag{2.5b}$$

we figure that the implicit Euler method succeeds, see again Figure 2.1.

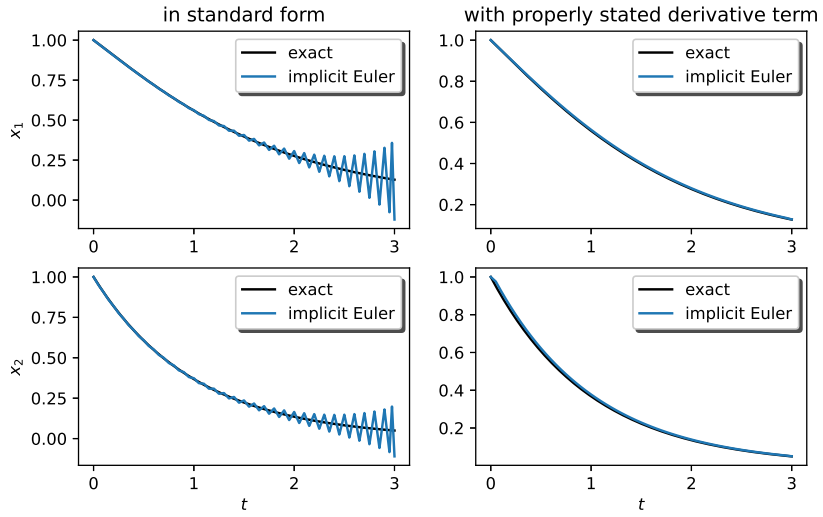


Figure 2.1: Numerical solution of the η-DAE in form (2.4) and (2.5) with $\eta = -0.52$ and step size $h = 0.5$.

However, it is not always given or achievable that a DAE meets the criterion of a properly stated derivative term. For the index analysis we weaken this requirement somewhat.

Definition 2.6 (semi-properly stated derivative term, [Jan15])
The DAE (2.2) has on $\mathcal{D} \times \mathcal{I}$ a semi-properly stated derivative term, if $\ker \frac{\partial}{\partial y} f(y, x, t)$ *and* $\operatorname{im} \frac{\partial}{\partial x} d(x, t)$ *are C^1-subspaces in $\mathbb{R}^n$, and the condition*

$$\operatorname{im} \frac{\partial}{\partial y} f(y, x, t) = \operatorname{im} \frac{\partial}{\partial y} f(y, x, t) \operatorname{im} \frac{\partial}{\partial x} d(x, t), \quad \forall (y, x, t) \in \mathbb{R}^n \times \mathcal{D} \times \mathcal{I},$$

holds.

The requirement of a semi-properly stated derivative term does not restrict DAEs in standard form at all.

Remark 2.7
A DAE in standard form (2.1) translated into one with nonlinear derivative by setting d as the identity has always a semi-properly stated derivative term.

2.2 Coupled Systems

Coupled systems can be understood as a collection of $r \geq 2$ so-called subsystems which share their unknown vector function, or parts thereof, with each other. One has to distinguish between two different origins of such coupled systems. First, there are the ones that arise from clearly separable problems, for instance different physical phenomena, which come along with their individual models and thus systems. If one assumes that these problems influence each other, a new holistic problem arises. The individual systems then become subsystems of the holistic problem. Then there is the second, opposing, origin of coupled problems where a single problem's deduced system is subdivided into or interpreted as multiple subsystems. Note that these philosophies can be combined as desired and it is usually a matter of tools and possibilities to treat the resulting subsystems. Whereas the first approach maintains the individual problems structure, and thus might be treated using well developed analysis, the second one is free of physical justification and might yield subsystems that are even easier to handle, see further the discussion in Chapter 6.3.

Let $\mathcal{I} \subset \mathbb{R}$ and $\mathcal{D} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_r \subset \mathbb{R}^{m_1+\cdots+m_r}$ be open subsets with $t_0 \in \mathcal{I}$. Consider the subsystems

$$f_i(\frac{\mathrm{d}}{\mathrm{d}t} d_i(x,t), x, t) = 0, \qquad \text{for } i = 1, \ldots, r \tag{2.6}$$

where $f_i : \mathbb{R}^{n_i} \times \mathcal{D} \times \mathcal{I} \to \mathbb{R}^{m_i}$ is continuous with existing partial derivatives of the first two arguments, and $d_i : \mathcal{D} \times \mathcal{I} \to \mathbb{R}^{n_i}$ is continuous differentiable. For each subsystem, we denote the intrinsic and complementary parts of the of the unknown vector function x with respectively

$$x_i \in \mathcal{D}_i \quad \text{and} \quad \bar{x}_i = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_r),$$

Definition 2.8 (coupled system)
We call (2.6) a coupled system if each subsystem, that is, for $i = 1, \ldots, r,$, the i-th equation under the prerequisite of known $\bar{x}_i = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_r)$

$$f_i(\frac{\mathrm{d}}{\mathrm{d}t} d_i(x_i, t; \bar{x}_i), x_i, t; \bar{x}_i) = 0$$

has a solution x_i, that is a continuous function on $\mathcal{I}$ with values in $\mathcal{D}_i$ so that $d_i(x_i(\cdot), \cdot; \bar{x}_i(\cdot))$ is continuous differentiable and x_i solves the equations pointwise on $\mathcal{I}$.

A coupled subsystem (2.6) itself is a system of the form (2.2) and represents either an implicit ODE or DAE with nonlinear derivative term. Further, (2.6) is often completed with the initial condition

$$x(t_0) = x_0 = (x_{0,1}, \ldots, x_{0,r}) \in \mathcal{D},$$

for an initial value $x_0 = (x_{0,1}, \ldots, x_{0,r}) \in \mathcal{D}$ which is assumed to be consistent.

Nevertheless, for our analysis it is sufficient to focus on additive coupled systems which will be introduced in the following.

Definition 2.9 (additive coupled)
A coupled system (2.6) is called additive coupled if its subsystems, for $i = 1, \dots, r$, comply the form

$$f_i(\frac{\mathrm{d}}{\mathrm{d}t} d_i(x_i, t), x_i, t) = c_i(\frac{\mathrm{d}}{\mathrm{d}t} \bar{d}_i(\bar{x}_i, t), \bar{x}_i, t) \tag{2.7}$$

with continuous $f_i : \mathbb{R}^{n_i} \times \mathcal{D}_i \times \mathcal{I} \to \mathbb{R}^{k_i}$ and $c_i : \mathbb{R}^{l_i} \times \bar{\mathcal{D}}_i \times \mathcal{I} \to \mathbb{R}^{k_i}$, which have existing partial derivatives of the first two arguments each, and continuous differentiable $d_i : \mathcal{D}_i \times \mathcal{I} \to \mathbb{R}^{n_i}$ and $\bar{d}_i : \bar{\mathcal{D}}_i \times \mathcal{I} \to \mathbb{R}^{l_i}$,

Note, these definitions apply to DAE and ODE subsystems. In order to specify the subsystems' classes we might say ODE-DAE additive coupled system in the case of $r = 2$. As for the overall systems (2.6) and (2.7), the can be represented by DAEs or (implicit) ODEs as well.

Example 2.10
For a given function $v : \mathbb{R} \to \mathbb{R}$, consider the two systems

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} y_1(t) + y_2(t) &= q \\ y_1(t) - v(t) &= 0 \end{aligned} \qquad \textit{and} \qquad \frac{\mathrm{d}}{\mathrm{d}t} z_1(t) = r$$

linked with each other by the right-hand sides $q := -z_1$ and $r := y_1$. Together, they form an additive coupled system

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} y_1(t) + y_2(t) &= -z_1(t) \\ y_1(t) - v(t) &= 0 \\ \frac{\mathrm{d}}{\mathrm{d}t} z_1(t) &= y_1(t) \end{aligned}$$

where the variables in notation of (2.6) read

$$x_1 = (y_1, y_2) \quad \textit{and} \quad x_2 = (z_1).$$

Considering the electric circuit modeling approach in Chapter 3, the coupling in Example 2.10 can be interpreted as adding a resistor parallel to the capacitor in a circuit which consists of a loop with a capacitor and a voltage source.

2.3 The Dissection Index

In the following we give a brief introduction to the dissection index concept with truncated projections developed by Lennart Jansen in [Jan15], albeit with slightly varying notation. This concept combines the advantages from both worlds, the projector based tractability index concept and the strangeness index concept.
The index, roughly speaking, can be understood as a tool of certain concepts to measure

how far away a DAE is to become an ODE, for instance the differentiation, perturbation, strangeness, tractability or dissection index concepts [BCP89; HLR89; KM94; LMT13; Jan15]. The necessity of these concepts becomes clear when trying to solve DAEs by means of numerical time integration in the sense that they are associated with several difficulties, see [HW91]. Depending on the effort involved, this obstacle can be circumvented, e.g. by decoupling the DAE to obtain the so-called inherent ODE with the help of constructive index concepts, or by reducing the index with stabilization methods, e. g. [Bau72; GLG85]. To give an example, it allows for a decoupling of circuit equations, see Chapter 3, with state independent projections.

The key point of the dissection index concept is given by the matrix chain strategy [LMT13] while using basis functions [KM06] which define the so-called kernel splitting pairs. Note that from here on the term matrix is understood to be exchangeable with the term matrix function.

Definition 2.11 (kernel splitting pair)
A pair of matrices $\{P, Q\}$ is called a kernel splitting pair of a given matrix M if $\operatorname{im} Q = \ker M$ *and the concatenation* $\begin{bmatrix} Q & P \end{bmatrix}$ *is nonsingular.*

Lemma 2.12 ([ST20])
Let $\{P, Q\}$ be a kernel splitting pair of a matrix M with full row rank. Then, MP is nonsingular.

Proof. It follows from the definition and full row rank of $M \begin{bmatrix} Q & P \end{bmatrix} = \begin{bmatrix} 0 & MP \end{bmatrix}$. □

Lemma 2.13 ([ST20])
Let M and N be two matrices with the same number of rows. Furthermore, assume that N is positive definite. If $\{P, Q\}$ is a kernel splitting pair of M then $\{V, Q\}$ with

$$V^\top := \begin{bmatrix} 0 & I_k \end{bmatrix} \begin{bmatrix} Q & P \end{bmatrix}^{-1} N \quad \text{and} \quad \operatorname{rank} I_k = k := \operatorname{rank} P$$

is also a kernel splitting pair of M.

Proof. By definition of a kernel splitting pair, we only have to show that $[Q\ V]$ is nonsingular. Assume that there exist x and y such that

$$Qx + Vy = 0. \tag{2.8}$$

For $S := \begin{bmatrix} 0 & I_k \end{bmatrix} \begin{bmatrix} Q & P \end{bmatrix}^{-1} = V^\top N^{-1}$, we obtain

$$\begin{aligned} SN^\top S^\top y = S(SN)^\top y = SVy = -SQx = -\begin{bmatrix} 0 & I_k \end{bmatrix} \begin{bmatrix} Q & P \end{bmatrix}^{-1} Qx \\ = -\begin{bmatrix} 0 & I_k \end{bmatrix} \begin{bmatrix} Q & P \end{bmatrix}^{-1} \begin{bmatrix} Q & P \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = -\begin{bmatrix} 0 & I_k \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = 0. \end{aligned}$$

Since $N^\top$ is positive definite, we can conclude $S^\top y = 0$. This implies $y = 0$ since S has full row rank. Regarding (2.8) and the full column rank of Q, we get also $x = 0$. □

Matrix Chain for Dissection Index Consider the DAE (2.2) with semi-properly stated derivative term. Let $(x^1, x, t) \in \mathbb{R}^n \times \mathcal{D} \times \mathcal{I}$ and

$$\begin{aligned} D(x^1, x, t) &:= \frac{\partial}{\partial x} d(x, t), \\ A(x^1, x, t) &:= \frac{\partial}{\partial y} f(d'(x, t), x, t), \\ B(x^1, x, t) &:= \frac{\partial}{\partial x} f(d'(x, t), x, t) \end{aligned}$$

where $(\cdot)'$ is the *jet-derivative operator*, see [LMT13], which introduces further jet variables x^1 as derivative placeholder. For the sake of human readability and conceptional understanding only, we omit the matrix functions' arguments from here on.

The *matrix chain*, associated with the dissection index, is initialized by choosing kernel splitting pairs $\{P, Q\}$ and $\{V, W\}$ of respectively AD and $(AD)^\top$ yielding

$$\begin{aligned} G_1 &:= V^\top ADP, & B_{x_1}^V &:= V^\top [BP + A(DP)'], & B_{y_1}^V &:= V^\top [BQ + A(DQ)'], \\ & & B_{x_1}^W &:= W^\top BP, & B_{y_1}^W &:= W^\top BQ. \end{aligned}$$

The chain continues for $i \geq 1$ by successively choosing kernel splitting pairs

- $\{P_{y_i}, Q_{y_i}\}$ and $\{V_{y_i}, W_{y_i}\}$ of respectively $B_{y_i}^W$ and its transposed
- $\{P_{x_i}, Q_{x_i}\}$ and $\{V_{x_i}, W_{x_i}\}$ of respectively $W_{y_i}^\top B_{x_i}^W$ and $(G_i Q_{x_i})^\top$

yielding, for correction term $R_i := B_{y_i}^V P_{y_i} (V_{y_i}^\top B_{y_i}^W P_{y_i})^{-1} V_{y_i}^\top B_{y_i}^W Q_{y_i}$,

$$\begin{aligned} G_{i+1} &:= V_{x_i}^\top G_i Q_{x_i}, & B_{x_{i+1}}^V &:= V_{x_i}^\top [B_{x_i}^V Q_{x_i} + G_i Q_{x_i}' - R_i], & B_{y_{i+1}}^V &:= V_{x_i}^\top B_{y_i}^V Q_{y_i}, \\ & & B_{x_{i+1}}^W &:= W_{x_i}^\top [B_{x_i}^V Q_{x_i} + G_i Q_{x_i}' - R_i], & B_{y_{i+1}}^W &:= W_{x_i}^\top B_{y_i}^V Q_{y_i}. \end{aligned}$$

The chain stops for smallest integer μ fulfilling $r_\mu = m$ with $r_0 := \operatorname{rank} AD$ and $r_i := r_{i-1} + \operatorname{rank} B_{y_i}^W$. Note, that past this criterion check, no further kernel splitting pairs have to be calculated.

Definition 2.14 (dissection index [Jan15])
Let the DAE (2.2) have a semi-properly stated derivative term, let f and d be sufficiently smooth and let $\mathcal{D} \times \mathcal{I}$ be further connected. Assuming that all the basis functions exist and have constant rank on their respective domain, the DAE (2.2) is said to be regular on $\mathcal{D} \times \mathcal{I}$ with dissection index μ if the matrix chain stops for μ.

The variables involved in the matrix chain are recursively defined, for $i = 1, \ldots, \mu$ by

$$x = Px_1 + Qy_1, \qquad x_{i-1} = P_{x_{i-1}} \tilde{x}_i + Q_{x_{i-1}} x_i, \qquad y_{i-1} = P_{y_{i-1}} \tilde{y}_i + Q_{y_{i-1}} y_i,$$

Remark 2.15
Given that the DAE (2.2) has even a properly stated derivative term implies $DQ = 0$ leading to the simplification $B_{y_1}^V = VBQ$ in the initial phase of the matrix chain, see [Jan15].

With the following lemma, an alternative end of the matrix chain is given, with which one can also determine the dissection index.

Lemma 2.16 ([Jan15])
A (2.2) has dissection index μ if and only if it has dissection index larger then $\mu - 1$ and $(W_y^*)^\top G_{\mu-1} Q_{x_{\mu-1}}$ is nonsingular whereby $\{V_y^*, W_y^*\}$ is a kernel splitting pair of $(B^V_{y_{\mu-1}} Q_{y_{\mu-1}})^\top$.

Remark
Note, that sometimes the term 'index 0' is used to describe ODEs even though, depending on the philosophy, ODEs are mostly clearly separated from DAEs. This has a practical reason, as we usually start our index analysis assuming a DAE (2.2) and figure out later that it was actually an (implicit) ODE.

2.4 Conclusions

With this chapter, the foundation for later analysis is laid. We introduced DAEs and pointed out their convenience when coupling different equation systems consisting of differential and/or algebraic equations. On the other hand, we have shown what structural difficulties come along with DAEs, concerning differentiability and consistent initialization, see Example 2.3. As for the differentiability, we introduced properly stated derivative terms for DAEs which often have a positive impact on time integration methods, see Example 2.5. Motivated by these structural difficulties, various index concepts were developed to measure them among which we introduced the dissection index concept developed in [Jan15] for later analysis of the here considered field/circuit coupled systems.

3 Electric Circuits

Witnessed in most homes across the globe, electric progress is a driving force in our society today. Both the large-scale electrotechnical behavior of electrical systems and the behavior of electronic components are typically described by means of electric circuits with lumped elements. This technique makes use of breaking down the considered system into a finite number of linked elements, which are assumed to behave in an idealized way. On the other hand, this is also how these systems are designed.

In order to simulate circuits within the framework of industrial applications, the modified nodal analysis (MNA), originally developed by Ho et. al., is one of the most frequently chosen approaches, cf. [HRB75; CL75; DK69; CDK87; FG05; RS09]. Others are for example the phaselock loop analysis or the sparse tableau analysis, see e. g. [Gar05; HBG71]. The MNA lead to the development of various circuit simulation softwares, for example `SPICE` (Simulation Program with Integrated Circuit Emphasis) [NP73] which also supports the modeling of more complex semiconductors elements, see for instance [MA87; Rei98; HN85] for a deeper insight.

In terms of transient simulation, the arising mathematical systems consist of mixed differential and algebraic equations, henceforth DAEs. Ever since then, DAEs were tried to be understood and analyzed from a numerical point of view, see e. g. [Gea71]. Because of the underlying graph structure, relationships between a circuit's topology and the underlying DAE's properties were explored, among which is in particular the DAE's index, see further [ET00; GF95]. However, analyzing the resulting DAE alone is not enough to keep pace with the challenging demands of the industry especially for the design and simulation of very-large-scale integration chips. Therefore, the solving process requires further ingredients such as *model order reduction*, see e. g. [SVR08; BS17; RS09], preferably going hand in hand without worsening the DAE's structure [Alì+13].

As for preparation of the coupled analysis in Chapter 5, we introduce a new circuit element, the *mock element*, as a placeholder for refined models such as the electromagnetic field device in Chapter 4. The name mock element is motivated by the *Functional Mock-up Interface*, a standardized formalism used to combine different models [Blo+11], such as from `Simulink` [Mat96]. This mock element approach, already used in [ST20], is similar to generalized element approach developed in [CDS19]. In a joint work, see [Cor+20], the idea of generalized elements was taken even further with focus on a systematic analysis and incorporation of the basic elements.
Together with the other common circuit elements, which are capacitors, resistors, inductors, voltage sources and current sources, we apply the MNA approach and decoupled the resulting DAE using the just introduced dissection index concept of [Jan15] introduced in Section 2. Note that henceforth the just mentioned elements such as capacitors, resistors and inductors

always refer to ideal ones and not the real devices. In Chapter 5, we exploit the mock element to provide various possibilities to incorporate an electromagnetic device whose model equations are introduced in Chapter 4. With this approach a convenient way is enabled to analyze the different coupling approaches especially in terms of convergence analysis when applying waveform relaxation methods to them, see Chapter 6. In addition to the mock element we also provide a new topological interpretation of certain matrix terms which arise during the decoupling process.

The chapter recalls and builds upon the results presented in [ST20, Section 2] and is therefore organized in a similar fashion as follows. First, we make the connection between electric circuits and graph theory which then allows us to derive a topological characterization guaranteeing a specific property of the deduced matrix terms. Followed by this, we introduce a lumped circuit model using the MNA approach with standard and mock elements. In preparation for Chapter 6, we finally provide a decoupled representation of the circuit's DAE using the dissection index concept introduced in Chapter 2.

3.1 Matrix Representation and Characterization of Circuit Structures

An electric circuit can be topologically interpreted as a graph. To be more precise, if considering multi-terminal elements such as transistors, the graph is a *hypergraph* $\mathcal{H} = (\mathcal{N}, \mathcal{E})$ which is to be understood as a generalization of graphs, see [Vol09]. Here, $\mathcal{N}$ denotes the set of nodes and $\mathcal{E}$ the set of hyperedges which can connect more then just two nodes.
Before providing a matrix representation of the circuit, we first reduce the possible hypergraph to an oriented graph whose edges allow only for two nodes. First, for each hyperedge we choose one incident node to be the reference node and introduce edges between this node to the non-reference nodes of the hyperedge. These edges are referred to as branches which are collected in $\mathcal{B}$. Next, we further assign an orientation to these branches which is directed towards their individual reference node. Finally, the oriented graph, given by $\mathcal{G} = (\mathcal{N}, \mathcal{B})$, is then a possible representation of the circuit. See Figure 3.1 for an example of how a multi-terminal circuit element may induce branches.

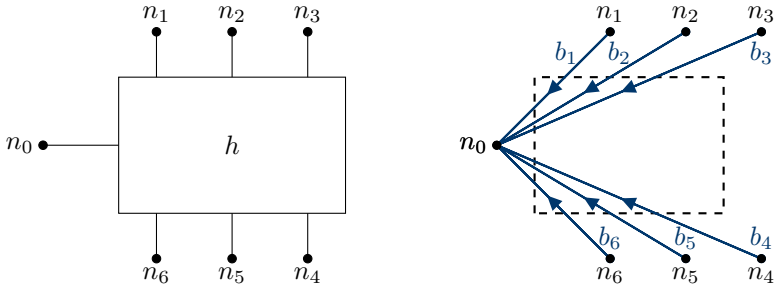


Figure 3.1: A multi-terminal element represented by a hyperedge $h \in \mathcal{H}$ with seven terminals $n_0, n_1, n_2, n_3, n_4, n_5, n_6 \in \mathcal{N}$. Choosing n_0 as the reference terminal yields six directed branches $b_1, b_2, b_3, b_4, b_5, b_6 \in \mathcal{B}$.

In this treatise we only consider *proper electric circuits* meaning that the deduced oriented graph $\mathcal{G} = (\mathcal{N}, \mathcal{B})$ meets the following requirements:

- $\mathcal{G}$ is connected.
- Each branch of $\mathcal{B}$ connects two different nodes of $\mathcal{N}$.
- Each node of $\mathcal{N}$ connects at least two branches of $\mathcal{B}$.

Notice that in graph theory $\mathcal{G}$ is often called a *multigraph* since it allows two nodes to be connected by an arbitrary number of branches. Further $\mathcal{G}$ is not allowed to have self-loops.

Similar to the circuit elements, we now choose one node, usually the ground node, to be the circuit's reference node. Without loss of generality, let the nodes and branches be labeled by ascending integers in their according sets and the circuit's reference node be the last node. With these preliminaries, we can now store the graph's structure, and hence the one of the circuit, by means of incidence matrices.

Definition 3.1 (incidence matrices)
For a circuit $\mathcal{G} = (\mathcal{N}, \mathcal{B})$, let $n \in \mathbb{N}$ be the number of nodes in $\mathcal{N}$ and $m \in \mathbb{N}$ the number of branches in $\mathcal{B}$.

(i) The matrix $A_{\text{full}} \in \{-1, 0, 1\}^{n \times m}$ defined by

$$(A_{\text{full}})_{i,j} := \begin{cases} 1 & \text{if node } i \text{ is a non-reference terminal of branch } j, \\ -1 & \text{if node } i \text{ is the reference terminal of branch } j, \\ 0 & \text{else} \end{cases}$$

is called full incidence matrix.

(ii) The matrix $A \in \{-1, 0, 1\}^{(n-1) \times m}$, obtained from A_{full} after eliminating the row corresponding to the predefined graph's reference node, is called reduced incidence matrix which we just call incidence matrix from here on.

Since a proper electric circuit is connected, its full incidence matrix has dependent rows contrary to its incidence matrix which motivated the latter introduction in the first place. Being familiar with graph theory's typical terminology, see A.2, A.3 and A.4, the following remark summarizes a few important properties taken from [DK69].

Remark 3.2
For a proper electric circuit it holds:

(i) The reduced incidence matrix has always full row rank.

(ii) Loops of the circuit correspond to linearly dependent columns of A and cutsets to linearly dependent rows of A.

(iii) Trees of the circuit form a nonsingular submatrix of A.

For decoupling the circuit equations in Section 3.3, we introduce the following notations for the sake of simplicity.

Definition 3.3 (cutset, loop, tree and path notation)
Let X and Y be disjoint branch types.

(i) An XY-cutset/-loop/-tree/-path is a cutset/loop/tree/path consisting of branches of types X and Y only.

(ii) Further type literals can be successively added to XY or deleted.

(iii) A $^+$ sign after a type literal indicates that the considered set consists of at least one branch of that specific type.

To give an example, an XY^+-loop is a loop consisting of branches of types X and Y but with at least one Y-type branch and a Y^+Z-cutset is a cutset consisting of branches of types Y and Z but at least with one Y-type branch.

The following theorem generalizes what in earlier work was only related to LI-cutsets and their incidence matrices to general element types with at least one specific and their consequences for the incidence matrix.

Let X, Y and Z be disjoint branch types with corresponding branch subsets $\mathcal{B}_\mathrm{X}$, $\mathcal{B}_\mathrm{Y}$ and $\mathcal{B}_\mathrm{Z} \subset \mathcal{B}$, respectively, fulfilling the relation $\mathcal{B}_\mathrm{Z} = \mathcal{B} \backslash (\mathcal{B}_\mathrm{X} \dot\cup \mathcal{B}_\mathrm{Y})$. Further, we denote the incidence matrices of the circuits represented by the subgraphs $(\mathcal{N}, \mathcal{B}_\mathrm{X})$, $(\mathcal{N}, \mathcal{B}_\mathrm{Y})$ and $(\mathcal{N}, \mathcal{B}_\mathrm{Z})$ with respectively A_X, A_Y and A_Z.

Theorem 3.4 (Theorem 1 from [ST20])
$\mathcal{G}$ has an XY^+-cutset, if and only if there exists an x such that $A_\mathrm{Z}^\top x = 0$ and $A_\mathrm{X}^\top x \neq 0$.

Proof. ($\Rightarrow$) Let $\mathcal{B}_c \subset \mathcal{B}_\mathrm{X} \cup \mathcal{B}_\mathrm{Y}$ form an XY^+-cutset of $\mathcal{G}$. The cutset divides $\mathcal{N}$ into two nonempty subsets $\mathcal{N}_1$ and $\mathcal{N}_2$ where all branches of $\mathcal{G}$ that connect nodes from $\mathcal{N}_1$ with nodes from $\mathcal{N}_2$ form the cutset $\mathcal{B}_c$. Without loss of generality, the reference node belongs to $\mathcal{N}_2$ and counted last as number n. We define $x \in \mathbb{R}^{n-1}$ as follows.

$$x_i := \begin{cases} 1 & \text{if node } i \text{ belongs to } \mathcal{N}_1, \\ 0 & \text{if node } i \text{ belongs to } \mathcal{N}_2. \end{cases}$$

Additionally, we introduce $x_n := 0$ for the reference node. Consider any $b_z \in \mathcal{B}_\mathrm{Z}$. It connects two nodes i_1 and i_2. Since $\mathcal{B}_c \subset \mathcal{B}_\mathrm{X} \cup \mathcal{B}_\mathrm{Y}$ we get $b_z \notin \mathcal{B}_c$. Consequently, the nodes i_1 and i_2 either belong both to $\mathcal{N}_1$ or both to $\mathcal{N}_2$ which yields $x_{i_1} = x_{i_2}$. By definition of the incidence matrix we obtain $A_\mathrm{Z}^\top x = 0$. Consider now any Y-type branch b_y from $\mathcal{B}_c$. It connects one node $i_1 \in \mathcal{N}_1$ with one node $i_2 \in \mathcal{N}_2$ yielding $x_{i_1} = 1$ and $x_{i_2} = 0$. For the corresponding column a_y of A_Y we obtain $a_y^\top x = a_{i_1,y} = \pm 1$. It results in $A_\mathrm{Y}^\top x \neq 0$.

($\Leftarrow$) Let $x \in \mathbb{R}^{n-1}$ such that $A_\mathrm{X}^\top x = 0$ and $A_\mathrm{X}^\top x \neq 0$. Since $A_\mathrm{X}^\top x \neq 0$, we know that there exists a column a_y of A_Y s. t. $a_y^\top x \neq 0$. We denote the branch corresponding to a_y by $b_y \in \mathcal{B}$. The branch b_y connects two nodes i_1 and i_2. Due to $a_y^\top x \neq 0$, we obtain that $x_{i_1} \neq 0$ or $x_{i_2} \neq 0$. We can assume that $x_{i_1} \neq 0$. Now we form a subset $\mathcal{N}_1$ of the nodes of $\mathcal{G}$ as follows: node $j \in \mathcal{N}$ belongs to $\mathcal{N}_1$ if and only if $x_j = x_{i_1}$. We define $\mathcal{N}_2 := \mathcal{N} \backslash \mathcal{N}_1$. Consequently, the node i_2 and the ground node belong to $\mathcal{N}_2$.

Let $\mathcal{B}_c$ be the set of all branches of $\mathcal{G}$ that connect nodes from $\mathcal{N}_1$ with nodes from $\mathcal{N}_2$. By construction, $\mathcal{B}_c$ forms a cutset of $\mathcal{G}$ with $b_y \in \mathcal{B}_c$. Consider any $b_z \in \mathcal{B}_\mathrm{Z}$. It connects two nodes j_1 and j_2. Let a_z be the corresponding branch of A_Z. Since $A_\mathrm{Z}^\top x = 0$, we get $a_z^\top x = 0$. If one of the nodes j_1 and j_2 is the ground node, say j_2, then $a_z^\top x = \pm x_{j_1}$ which yields $x_{j_1} = 0$. Consequently, j_1 and j_2 belong to $\mathcal{N}_2$. If both nodes j_1 and j_2 are different from the ground node then $a_z^\top x = \pm(x_{j_1} - x_{j_2})$ which implies $x_{j_1} = x_{j_2}$. It means, that the nodes j_1 and j_2 belong either both to $\mathcal{N}_1$ or both to $\mathcal{N}_2$. Therefore, $b_z \notin \mathcal{B}_c$ and we obtain that $\mathcal{B}_c$ is a XY^+-cutset. □

With the help of Theorem 3.4 we derive a different topological interpretation of the expression $Q_\mathrm{Z}^\top A_\mathrm{Y}$ being non-trivial. Usually, $Q_\mathrm{Z}^\top A_\mathrm{Y} = 0$ is equivalent to the existence of Z-paths between end nodes of Y-branches, see e. g. [ET00].

Corollary 3.5 (Corollary 1 from [ST20])
Let the columns of Q_Z form a basis of the kernel of $A_\mathrm{Z}^\top$ that means $\operatorname{im} Q_\mathrm{Z} = \ker A_\mathrm{Z}^\top$. Then, $\mathcal{G}$ has an XY^+-cutset if and only if $Q_\mathrm{Z}^\top A_\mathrm{Y} \neq 0$.

Proof. $Q_\mathrm{Z}^\top A_\mathrm{Y} \neq 0$ is equivalent to $A_\mathrm{X}^\top Q_\mathrm{Z} \neq 0$. This, furthermore, is equivalent to the existence of an $x \in \operatorname{im} Q_\mathrm{Z}$ with $A_\mathrm{X}^\top x \neq 0$. As $A_\mathrm{Z}^\top x = 0$ if and only if $x \in \operatorname{im} Q_\mathrm{Z}$, we deduce the assertion to be fulfilled by Theorem 3.4. □

For the circuits under consideration here, Corollary 3.5 implies, for example, that nodes of mock element branches are connected via CVR-paths if there is no LIM^+-cutset. That follows immediately by using the type substitutions $X = LI$, $Y = M$ and $Z = CVR$. We conclude this section with a last Lemma.

Lemma 3.6 (Lemma 1 from [ST20])
Let the columns of Q_X and Q_Y form bases of the kernels of $A_\mathrm{X}^\top$ and $A_\mathrm{X}^\top Q_\mathrm{X}$, respectively. Further, let the columns of Q_XY be as basis of the kernel of the concatenated matrix $\begin{bmatrix} A_\mathrm{X} & A_\mathrm{Y} \end{bmatrix}^\top$. Then, $\ker Q_\mathrm{XY}^\top = \ker Q_\mathrm{Y}^\top Q_\mathrm{X}^\top$.

Proof. The assertion follows from

$$z \in \ker Q_\mathrm{Y}^\top Q_\mathrm{X}^\top \quad \Leftrightarrow \quad Q_\mathrm{Y}^\top Q_\mathrm{X}^\top z = 0 \quad \Leftrightarrow \quad Q_\mathrm{X}^\top z \in \operatorname{im} Q_\mathrm{X}^\top A_\mathrm{Y}$$

and

$$\begin{aligned} z \in \ker Q_\mathrm{XY}^\top \quad &\Leftrightarrow \quad z \in \operatorname{im} \begin{bmatrix} A_\mathrm{X} & A_\mathrm{Y} \end{bmatrix} \quad \Leftrightarrow \quad \exists w \exists v : \; z = A_\mathrm{Y} w + A_\mathrm{X} v \\ &\Leftrightarrow \quad \exists w : \; z - A_\mathrm{Y} w \in \operatorname{im} A_\mathrm{X} \quad \Leftrightarrow \quad \exists w : \; Q_\mathrm{X}^\top (z - A_\mathrm{Y} w) = 0. \end{aligned}$$

□

3.2 Lumped Circuit Modeling

A typical way to model electric circuits is my means of lumped element models which simplify the elements' description of behavior reduced to their essential and often idealistic function in the circuit. Note that as a consequence, especially spatial effects are neglected. These lumped element models are then combined with Kirchhoff's circuit laws yielding a sufficient set of equations which are usually organized my means of the MNA approach or dialects thereof, e. g the conventional or flux/charge oriented one, see further [CDK87; GF99]. Focusing on the conventional MNA, the quantities of interest are the voltage drops and currents across the elements and potentials at their joints.

Let $\mathcal{G} = (\mathcal{N}, \mathcal{B})$ be the circuit's graph representation as of Section 3.1 with $n \in \mathbb{N}$ nodes and $m \in \mathbb{N}$ branches whereby again without loss of generality the graph's reference node's label is n. In accordance with the just introduced terminology, for some time interval $\mathcal{I} := [t_0, T] \subset \mathbb{R}$ we introduce the vector-functions $\mathbf{i}, \mathbf{v} : \mathcal{I} \to \mathbb{R}^m$ and $\mathbf{e}_{\text{full}} : \mathcal{I} \to \mathbb{R}^n$ representing all *branch currents*, *branch voltages* and *node potentials*, respectively. In addition to that, with $\mathbf{e} : \mathcal{I} \to \mathbb{R}^{n-1}$ we denote the node potentials reduced by the graphs reference node. The circuit quantities with their SI units can be fount in Table 3.1 For the sake of readability we omit the time argument of unknown quantities in the continuing part of the chapter.

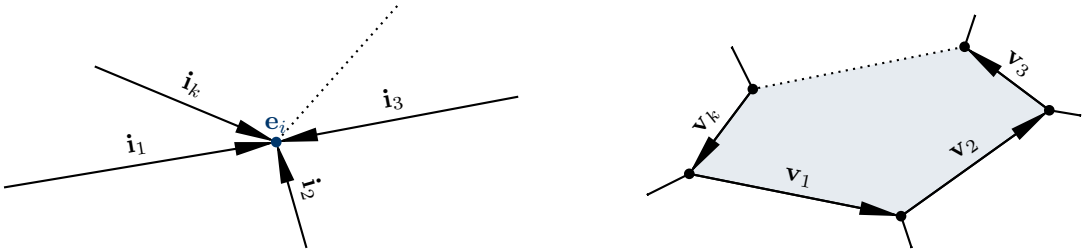
notation	SI unit	quantity
$\mathbf{i}$	A	branch currents
$\mathbf{v}$	V	branch voltages
$\mathbf{e}$	V	node potentials

Table 3.1: Electric circuit quantities and their SI units.

Kirchhoff's Circuit Laws The very foundation of circuit equations for lumped circuits are given by Kirchhoff's circuit laws, see e. g. [DK69], which are

Kirchhoff's current law (KCL) : For any node and at any time the algebraic sum of all branch currents entering or leaving the node is zero, see Figure 3.2a.

Kirchhoff's voltage law (KVL) : For any loop and at any time the algebraic sum of all branch voltages around the loop is zero, see Figure 3.2b.



(a) Kirchhoff's current law

(b) Kirchhoff's voltage law

Figure 3.2: Illustration of Kirchhoff's circuit laws.

For A being the circuits incidence matrix as of Definition 3.1, then Kirchhoff's current and voltage law yield

$$A\mathbf{i} = 0 \tag{3.1a}$$

$$\mathbf{v} = A^\top \mathbf{e}. \tag{3.1b}$$

given that the potentials ambiguity is fixed by setting the circuit's reference node's potential to zero.

Constitutive Element Equations Constitutive element equations provide the relations between branch currents and voltages. Together with Kirchhoff's circuit laws they allow for the formulation of a determinate system. For an arbitrary circuit element consider the dissection $\mathbf{i} = (\mathbf{i}_{\text{elem}}, \mathbf{i}_{\text{compl}})$ and $\mathbf{v} = (\mathbf{v}_{\text{elem}}, \mathbf{v}_{\text{compl}})$ according to those branches which belong to this specific element and those which do not. Usually, the constitutive element equation takes one out of two forms. To be more precise,

- the circuit element is called current controlling if the constitutive equation explicitly determines the branch currents, i. e.

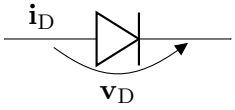
$$\mathbf{i}_{\text{elem}} = f_{\text{elem}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\text{elem}}(\mathbf{i}_{\text{compl}}, \mathbf{v}, t), \mathbf{i}_{\text{compl}}, \mathbf{v}, t); \tag{3.2}$$

- the circuit element is called voltage controlling if the constitutive equation explicitly determines the branch voltages, i. e.

$$\mathbf{v}_{\text{elem}} = f_{\text{elem}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\text{elem}}(\mathbf{i}, \mathbf{v}_{\text{compl}}, t), \mathbf{i}, \mathbf{v}_{\text{compl}}, t), \tag{3.3}$$

for some functions f_{elem} and d_{elem}. The forms (3.2) and (3.3) of constitutive equations cover the basic types of elements and even more. The basic *two-terminal elements* that are considered in this treatise are found in Table 3.2. Note that these elements are possibly nonlinear which makes, for instance, the `SPICE` diode model given by

$$\mathbf{i}_{\mathrm{D}} = I_S \left(e^{\frac{\mathbf{v}_{\mathrm{D}}}{N V_T}} - 1 \right) \tag{3.4}$$



a special case of the resistor. The equation (3.4) is called *Shockley equation*, with diode voltage and current $\mathbf{v}_\mathrm{D}$ and $\mathbf{i}_\mathrm{D}$, respectively. The term $V_T := \frac{k\bar{T}}{q}$ refers to the *thermal voltage* using constants

$$k = 1.3806226 \cdot 10^{-23} \qquad \text{Boltzmann constant in } \left[\frac{\mathrm{J}}{\mathrm{K}}\right],$$
$$q = 1.6021918 \cdot 10^{-19} \qquad \text{elementary charge in [C]}$$

and the parameters I_S, N and $\bar{T}$ are respectively the *reverse biased saturation current*, the *ideality factor* and the *nominal temperature*.

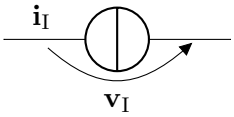
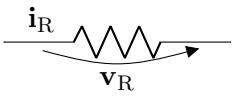
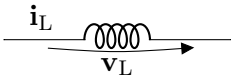
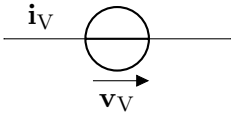
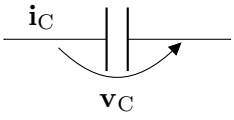
name	constitutive equation	symbol
current source	$\mathbf{i}_\mathrm{I} = i_s(t)$	
resistor	$\mathbf{i}_\mathrm{R} = g_\mathrm{R}(\mathbf{v}_\mathrm{R}, t)$	
inductor	$\mathbf{i}_\mathrm{L} = \frac{\mathrm{d}}{\mathrm{d}t}\phi_\mathrm{L}(\mathbf{v}_\mathrm{L}, t)$	
voltage source	$\mathbf{v}_\mathrm{V} = v_s(t)$	
capacitor	$\mathbf{v}_\mathrm{C} = \frac{\mathrm{d}}{\mathrm{d}t}q_\mathrm{C}(\mathbf{i}_\mathrm{C}, t)$	

Table 3.2: Basic two-terminal circuit elements with their constitutive element equations and symbols. Current and voltage quantities as well as the source functions v_s and i_s and characteristic functions $g_\mathrm{R}, \phi_\mathrm{L}$ and q_C are scalar valued if there is only one branch their type.

Following the approach of [ST20] in a similar way, we introduce a new current controlling element for the later inclusion of refined models, the mock element

$$\mathbf{i}_\mathrm{M} = f_\mathrm{M}(\frac{\mathrm{d}}{\mathrm{d}t}d_\mathrm{M}(\mathbf{v}_\mathrm{M}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}_\mathrm{M}, \mathbf{v}_\mathrm{M}, \mathbf{u}_\mathrm{M}) \tag{3.5}$$

where $\mathbf{u}_\mathrm{M}$ fulfills the implicit ODE

$$M_\mathrm{M}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}_\mathrm{M} + b_\mathrm{M}(\mathbf{u}_\mathrm{M}) = c_\mathrm{M}(\frac{\mathrm{d}}{\mathrm{d}t}d_\mathrm{M}(\mathbf{v}_\mathrm{M}), \mathbf{v}_\mathrm{M}) \tag{3.6}$$

instead of an explicit one. Contrary to the basic two-terminal elements in Table 3.2, we may expect multiple terminals and hence the characteristic function f_M and its derivative term d_M are likely to be vector-valued as well as $\mathbf{i}_\mathrm{M}$ and $\mathbf{v}_\mathrm{M}$. The mock element's intrinsic equation (3.6), for nonsingular matrix M_M, function b_M and coupling function c_M can be of any dimension, but always determines the intrinsic variable $\mathbf{u}_\mathrm{M}$. Note that a mock element can represent a variety of elements, even equivalent circuits of transistors and diodes given that their capacitive contribution is not neglected in the modeling. To give an example, consider the MOSFET equivalent circuit of Figure 3.3 as part of a NAND-gate model.

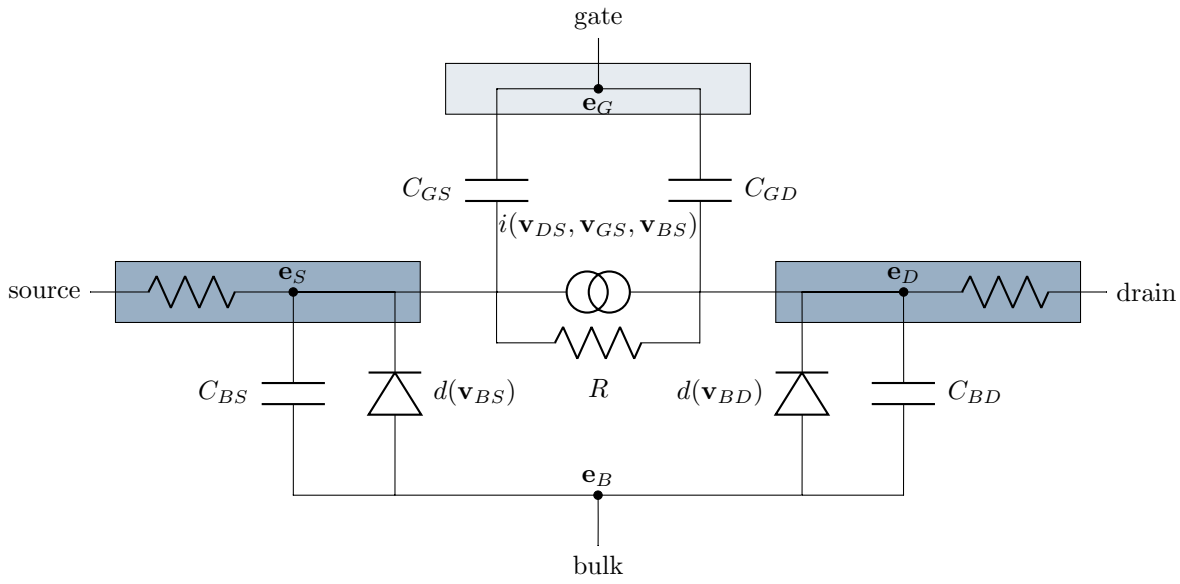


Figure 3.3: MOSFET equivalent circuit as used in [FG94] and [Gün95]. All circuit nodes $\mathbf{e}_S, \mathbf{e}_D, \mathbf{e}_G$ and $\mathbf{e}_B$ are connected by a capacitive path.

Modified Nodal Analysis A complete lumped circuit model for transient simulation can be obtained following the MNA approach. To understand its general concept, we dissect the branch quantities such that $\mathbf{i} = (\mathbf{i}_\mathrm{cur}, \mathbf{i}_\mathrm{vol})$, $\mathbf{v} = (\mathbf{v}_\mathrm{cur}, \mathbf{v}_\mathrm{vol})$, in accordance with aboves classifications (3.2) and (3.3). In a similar fashion we proceed with the incidence matrix, i. e. $A = \begin{bmatrix} A_\mathrm{cur} & A_\mathrm{vol} \end{bmatrix}$ complying with

$$A\mathbf{i} = A_\mathrm{cur}\mathbf{i}_\mathrm{cur} + A_\mathrm{vol}\mathbf{i}_\mathrm{vol} \quad \text{and} \quad \begin{pmatrix} \mathbf{v}_\mathrm{cur} \\ \mathbf{v}_\mathrm{vol} \end{pmatrix} = \begin{bmatrix} A_\mathrm{cur}^\top \\ A_\mathrm{vol}^\top \end{bmatrix} \mathbf{e}.$$

Following the same sorting as the quantities, we concatenate all the current and voltage controlling constitutive equations into respectively f_cur, d_cur and f_vol, d_vol. Motivated by the empirical observation that a circuit usually contains more branches than nodes, the branch voltages are

replaced by terms of potentials, that is exploiting KVL by inserting (3.1b) where possible. In order to get rid of even more quantities, we insert further f_{cur} into KCL (3.1a) with regards to the just introduced splitting and obtain from the circuits laws (3.1):

$$A_{\mathrm{cur}} f_{\mathrm{cur}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{cur}}(\mathbf{i}, A^\top \mathbf{e}, t), \mathbf{i}, A^\top \mathbf{e}, t) + A_{\mathrm{vol}} \mathbf{i}_{\mathrm{vol}} = 0, \tag{3.7a}$$

$$f_{\mathrm{vol}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{vol}}(\mathbf{i}, A^\top \mathbf{e}, t), \mathbf{i}, A^\top \mathbf{e}, t) - A_{\mathrm{vol}}^\top \mathbf{e} = 0. \tag{3.7b}$$

The system (3.7) is called the *modified nodal analysis* (MNA) in its conventional formulation and represents a DAE. In application these equations are generated automatically from net lists providing the node to branch element relations and the elements' characteristic functions. Note that using only the basic elements from Table 3.2, the resulting MNA system's unknowns are the nodal potentials $\mathbf{e}$ and the currents $\mathbf{i}_{\mathrm{compl}}$ of the voltage controlling elements only, see [ET00].

Assumption 3.7
Let the circuit consist of capacitors (C), resistors (R), inductors (L), voltage sources (V), current sources (I) and mock elements (M) only.

Given Assumption 3.7, we denote with $m_{\mathrm{C}}, m_{\mathrm{R}}, m_{\mathrm{L}}, m_{\mathrm{V}}, m_{\mathrm{I}}$ and $m_{\mathrm{M}} \in \mathbb{N}$ be the number of branches induced by the element types regarding the subscript. we can sort the currents and the columns of the incidence matrix according to the element types by introducing

$$\begin{aligned}
\mathbf{i} &= (\mathbf{i}_{\mathrm{C}}, \mathbf{i}_{\mathrm{R}}, \mathbf{i}_{\mathrm{L}}, \mathbf{i}_{\mathrm{V}}, \mathbf{i}_{\mathrm{I}}, \mathbf{i}_{\mathrm{M}}) : \mathcal{I} \to \mathbb{R}^{m_{\mathrm{C}}+m_{\mathrm{R}}+m_{\mathrm{L}}+m_{\mathrm{V}}+m_{\mathrm{I}}+m_{\mathrm{M}}}, \\
A &= \begin{bmatrix} A_{\mathrm{C}} & A_{\mathrm{R}} & A_{\mathrm{L}} & A_{\mathrm{V}} & A_{\mathrm{I}} & A_{\mathrm{M}} \end{bmatrix} \in \{-1, 0, 1\}^{(n-1)\times(m_{\mathrm{C}}+m_{\mathrm{R}}+m_{\mathrm{L}}+m_{\mathrm{V}}+m_{\mathrm{I}}+m_{\mathrm{M}})}.
\end{aligned}$$

With $q_{\mathrm{C}} : \mathbb{R}^{m_{\mathrm{C}}} \times \mathcal{I} \to \mathbb{R}^{m_{\mathrm{C}}}, g_{\mathrm{R}} : \mathbb{R}^{m_{\mathrm{R}}} \times \mathcal{I} \to \mathbb{R}^{m_{\mathrm{R}}}, \phi_{\mathrm{L}} : \mathbb{R}^{m_{\mathrm{L}}} \times \mathcal{I} \to \mathbb{R}^{m_{\mathrm{L}}}, i_s : \mathcal{I} \to \mathbb{R}^{m_{\mathrm{I}}}$ and $v_s : \mathcal{I} \to \mathbb{R}^{m_{\mathrm{V}}}$ we describe the element type-wise characteristic functions, resulting from concatenation of the ones in Table 3.2. Then, for $\mathbf{x} = (\mathbf{e}, \mathbf{i}_{\mathrm{L}}, \mathbf{i}_{\mathrm{V}})$ and $\mathbf{w} = (\mathbf{w}_{\mathrm{C}}, \mathbf{w}_{\mathrm{L}})$ the MNA system (3.7) takes the following special form

$$f_{\mathrm{MNA}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA}}(\mathbf{x}, t), \mathbf{x}, t) = c_{\mathrm{MNA}}(\mathbf{i}_{\mathrm{M}}) \tag{3.8}$$

where

$$f_{\mathrm{MNA}}(\mathbf{w}, \mathbf{x}, t) = \begin{pmatrix} A_{\mathrm{C}} \mathbf{w}_{\mathrm{C}} + A_{\mathrm{R}} g_{\mathrm{R}}(A_{\mathrm{R}}^\top \mathbf{e}, t) + A_{\mathrm{L}} \mathbf{i}_{\mathrm{L}} + A_{\mathrm{V}} \mathbf{i}_{\mathrm{V}} + A_{\mathrm{I}} i_s(t) \\ \mathbf{w}_{\mathrm{L}} - A_{\mathrm{L}}^\top \mathbf{e} \\ v_s(t) - A_{\mathrm{V}}^\top \mathbf{e} \end{pmatrix},$$

$$d_{\mathrm{MNA}}(\mathbf{x}, t) = \begin{pmatrix} q_{\mathrm{C}}(A_{\mathrm{C}}^\top \mathbf{e}, t) \\ \phi_{\mathrm{L}}(\mathbf{i}_{\mathrm{L}}, t) \end{pmatrix}, \qquad c_{\mathrm{MNA}}(\mathbf{i}_{\mathrm{M}}) = \begin{pmatrix} -A_{\mathrm{M}} \mathbf{i}_{\mathrm{M}} \\ 0 \\ 0 \end{pmatrix}$$

and $\mathbf{i}_{\mathrm{M}}$ given by a possible concatenation of mock element characteristic functions (3.5), denoted likewise. Note that also in the mock elements' intrinsic function is then given as a concatenation of multiple equations (3.6). Following MNA approach, we further replace $\mathbf{v}_{\mathrm{M}} = A_{\mathrm{M}}^\top \mathbf{e}$ in the mock elements' description (3.5) and (3.6).

In the following we collect some assumptions which are usually fulfilled in connection with the classical MNA, that is without mock elements. These assumptions are considered as standard and are not major limitations but mainly reflect compliance with physical properties, see e. g. [Fos92; ET00].

Assumption 3.8 (standard circuit assumptions)

Passivity *All resistors, inductors and capacitors in the circuit show a passive behavior, i. e.* $q_{\mathrm{C}}, g_{\mathrm{R}}$ *and* ϕ_{L} *are strongly monotone.*

Consistency *The circuit contains neither* V*-loops nor* I*-cutsets.*

Smoothness *The characteristic functions* g_{R}*,* q_{C} *and* ϕ_{L} *are globally Lipschitz continuous and the latter two are additionally continuously differentiable. The source functions* i_s *and* v_s *are twice continuously differentiable.*

First, with the passivity property we guarantee that the circuit elements consume energy and do not produce energy. Secondly, the consistency prevents the circuit from having a shortcut. It is a necessarily assumption for the existence of a (unique) solution, see [ET00]. Finally, with we assume sufficient smoothness of the characteristic element functions which are further required to guarantee existence and uniqueness of solutions.

In order to obtain a unique solution for the DAE (3.8), under the prerequisite of known and sufficient smooth $\mathbf{i}_{\mathrm{M}}$, one usually introduces an additional initial condition

$$x_0 = \mathbf{x}(t_0),$$

for some consistent initial value $x_0 \in \mathbb{R}^{n-1+m_{\mathrm{L}}+m_{\mathrm{V}}}$, cf. [Est00; ET00; Est02; Mat12].

Example 3.9 (electric circuit)
Consider the circuit given in Figure 3.4. For the orientation contained therein, we obtain an incidence matrix by

$$A = \begin{bmatrix} A_{\mathrm{C}} & A_{\mathrm{R}} & A_{\mathrm{L}} & A_{\mathrm{V}} & A_{\mathrm{I}} & A_{\mathrm{M}} \end{bmatrix} = \left(\begin{array}{c|cc|c|c|c} 0 & 1 & 0 & 0 & -1 & 0 \\ -1 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{array}\right).$$

For $\mathbf{x} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \mathbf{i}_{\mathrm{L}}, \mathbf{i}_{\mathrm{V}})$*, the MNA circuit equations, which form the DAE (3.8), read*

$$\begin{aligned} R(\mathbf{e}_1 - \mathbf{e}_2) - \mathbf{i}_{\mathrm{V}} &= 0, \\ \frac{\mathrm{d}}{\mathrm{d}t}(C\mathbf{e}_2) - R(\mathbf{e}_1 - \mathbf{e}_2) - i_s &= 0, \\ \mathbf{i}_{\mathrm{L}} + i_s &= 0, \\ d(\mathbf{e}_4) - \mathbf{i}_{\mathrm{L}} &= 0, \\ \frac{\mathrm{d}}{\mathrm{d}t}(L\mathbf{i}_{\mathrm{L}}) - (\mathbf{e}_3 - \mathbf{e}_4) &= 0, \\ v_s + \mathbf{e}_1 &= 0 \end{aligned}$$

with given parameters C, R, L, diode function d, see e. g. (3.4) and source functions i_s and v_s.

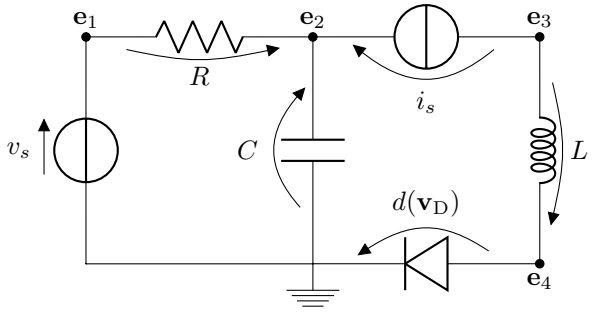


Figure 3.4: Exemplary circuit containing elements complying with Assumption 3.7.

3.3 Decoupling of the Lumped Circuit DAE

In this section we provide various decoupled formulations of (3.8) into the inherent ODE and the algebraic constraints depending on further topological assumptions. Their decoupling utilizes the dissection index concept of [Jan15] with truncated projections introduced in Chapter 2 and is similar to the projector based decoupling used in [ET00].

Because the mock element from our point of view is still a black box, we have to meet an additional assumption which is required only by the following decoupling theorems and not to be understood as a physical restriction.

Assumption 3.10 (No IM-cutsets)
The circuit contains no IM-cutsets.

For the sake of simplicity we define some new terms

$$s_c(\mathbf{i}_\mathrm{M}) := A_\mathrm{M}\mathbf{i}_\mathrm{M}, \qquad s_i(t) := A_\mathrm{I} i_s(t), \qquad s_v(t) := v_s(t).$$

Further, from hereon we drop the time dependency as an argument from all terms.

Theorem 3.11 (Theorem 2 from [ST20])
If Assumptions 3.7, 3.8 and 3.10 are satisfied, then there exist functions f_0, f_1, f_2 that are at least locally Lipschitz continuous as well as a (constant) matrix M_3, a nonsingular bounded matrix function M_1 and a (constant) nonsingular transformation matrix $T = \begin{bmatrix} T_0 & T_1 & T_2 & T_3 \end{bmatrix}$ such that the DAE (3.8) can be globally decoupled into an equivalent system of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = f_0(\mathbf{y}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c(\mathbf{i}_\mathrm{M})), \tag{3.9a}$$

$$\mathbf{z}_1 = M_1(\mathbf{y}, \mathbf{z}_3)\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3 + f_1(\mathbf{y}, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c(\mathbf{i}_\mathrm{M})), \tag{3.9b}$$

$$\mathbf{z}_2 = f_2(\mathbf{y}, \mathbf{z}_3, s_v, s_i, s_c(\mathbf{i}_\mathrm{M})), \tag{3.9c}$$

$$\mathbf{z}_3 = M_3\begin{pmatrix} s_i + s_c(\mathbf{i}_\mathrm{M}) \\ s_v \end{pmatrix} \tag{3.9d}$$

in that, for given $\mathbf{i}_\mathrm{E} \in C^1$*, the function* $\mathbf{x}$ *is a solution of* (3.8) *with* $\mathbf{x}(t_0) = x_0$ *if and only if* $\bar{\mathbf{x}}$*, defined by* $\mathbf{x} = T\bar{\mathbf{x}} = T_0\mathbf{y} + T_1\mathbf{z}_1 + T_2\mathbf{z}_2 + T_3\mathbf{z}_3$*, is a function solving the decoupled system* (3.9) *on* $\mathcal{I}$ *with* $\mathbf{y}(t_0) = y_0$ *satisfying*

$$x_0 = T_0 y_0 + T_1\mathbf{z}_1(t_0) + T_2\mathbf{z}_2(t_0) + T_3\mathbf{z}_3(t_0).$$

Proof. The proof combines the idea of the dissection concept [Jan15] for DAEs with the projector based decoupling [LMT13; ET00] for circuit DAEs. Due to Assumption 3.7, the DAE (3.8) reads

$$A_\mathrm{C}\frac{\mathrm{d}}{\mathrm{d}t}q_\mathrm{C}(A_\mathrm{C}^\top\mathbf{e}) + A_\mathrm{R}g_\mathrm{R}(A_\mathrm{R}^\top\mathbf{e}) + A_\mathrm{L}\mathbf{i}_\mathrm{L} + A_\mathrm{V}\mathbf{i}_\mathrm{V} + s_i + s_c(\mathbf{i}_\mathrm{M}) = 0, \tag{3.10a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_\mathrm{L}(\mathbf{i}_\mathrm{L}) - A_\mathrm{L}^\top\mathbf{e} = 0, \tag{3.10b}$$

$$-A_\mathrm{V}^\top\mathbf{e} + s_v = 0. \tag{3.10c}$$

Let $\{P, Q\}$ be kernel splitting pairs of the following matrices M:

$\{P,Q\}$	$\{P_\mathrm{C},Q_\mathrm{C}\}$	$\{P_\mathrm{V},Q_\mathrm{V}\}$	$\{P_\mathrm{R},Q_\mathrm{R}\}$	$\{\bar{P}_\mathrm{V},\bar{Q}_\mathrm{V}\}$	$\{\bar{P}_\mathrm{L},\bar{Q}_\mathrm{L}\}$	$\{P_\mathrm{e},Q_\mathrm{e}\}$
M	$A_\mathrm{C}^\top$	$A_\mathrm{V}^\top Q_\mathrm{C}$	$A_\mathrm{R}^\top Q_\mathrm{C}Q_\mathrm{V}$	$P_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{V}$	$Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{L}$	$\bar{Q}_\mathrm{V}^\top A_\mathrm{V}^\top P_\mathrm{C}$

We use them to split $\mathbf{e}$, $\mathbf{i}_\mathrm{L}$ and $\mathbf{i}_\mathrm{V}$ as follows:

$$\mathbf{e} = Q_\mathrm{C}\big[Q_\mathrm{V}(Q_\mathrm{R}\mathbf{z}_{1l} + P_\mathrm{R}\mathbf{z}_{2r}) + P_\mathrm{V}\mathbf{z}_{2v}\big] + P_\mathrm{C}[Q_\mathrm{e}\mathbf{y}_e + P_\mathrm{e}\mathbf{z}_{3e}]\,, \tag{3.11a}$$

$$\mathbf{i}_\mathrm{L} = \bar{Q}_\mathrm{L}\bar{\mathbf{y}}_l + \bar{P}_\mathrm{L}\bar{\mathbf{z}}_{3l}\,, \tag{3.11b}$$

$$\mathbf{i}_\mathrm{V} = \bar{Q}_\mathrm{V}\bar{\mathbf{z}}_{1v} + \bar{P}_\mathrm{V}\bar{\mathbf{z}}_{2v} \tag{3.11c}$$

and collect the new variables as $\mathbf{y} := (\mathbf{y}_e, \bar{\mathbf{y}}_l)$, $\mathbf{z}_1 := (\mathbf{z}_{1l}, \bar{\mathbf{z}}_{1v})$, $\mathbf{z}_2 := (\mathbf{z}_{2v}, \mathbf{z}_{2r}, \bar{\mathbf{z}}_{2v})$, $\mathbf{z}_3 := (\mathbf{z}_{3e}, \bar{\mathbf{z}}_{3l})$. Exploiting the splitting pairs' properties we introduce

$$\begin{aligned}
\hat{g}_\mathrm{R}(\mathbf{z}_{2r}, \mathbf{z}_{2v}, \mathbf{y}_e, \mathbf{z}_{3e}) &:= g_\mathrm{R}(A_\mathrm{R}^\top[Q_\mathrm{C}(Q_\mathrm{V}P_\mathrm{R}\mathbf{z}_{2r} + P_\mathrm{V}\mathbf{z}_{2v}) + P_\mathrm{C}(Q_\mathrm{e}\mathbf{y}_e + P_\mathrm{e}\mathbf{z}_{3e})]) \\
&= g_\mathrm{R}(A_\mathrm{R}^\top\mathbf{e}), \\
\hat{C}(\mathbf{y}_e, \mathbf{z}_{3e}) &:= P_\mathrm{C}^\top A_\mathrm{C} q_\mathrm{C}'(A_\mathrm{C}^\top P_\mathrm{C}[Q_\mathrm{e}\mathbf{y}_e + P_\mathrm{e}\mathbf{z}_{3e}])A_\mathrm{C}^\top P_\mathrm{C} = P_\mathrm{C}^\top A_\mathrm{C} q_\mathrm{C}'(A_\mathrm{C}^\top\mathbf{e})A_\mathrm{C}^\top P_\mathrm{C}, \\
\hat{L}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l}) &:= \phi_\mathrm{L}'(\bar{Q}_\mathrm{L}\bar{\mathbf{y}}_l + \bar{P}_\mathrm{L}\bar{\mathbf{z}}_{3l}) = \phi_\mathrm{L}'(\mathbf{i}_\mathrm{L}).
\end{aligned}$$

The function $\hat{g}_\mathrm{R}$ is globally Lipschitz continuous since g_R is globally Lipschitz continuous. Furthermore, $\hat{C}(\mathbf{y}_e, \mathbf{z}_{3e})$ and $\hat{L}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})$ are positive definite since $A_\mathrm{C}^\top P_\mathrm{C}$ has full column rank and q_C as well as ϕ_L are strongly monotone.

We proceed with some additional splitting pairs to split the equations. We choose kernel splitting pairs $\{V_\mathrm{C}(\mathbf{y}_e, \mathbf{z}_{3e}), Q_\mathrm{e}\}$ and $\{\bar{V}_\mathrm{L}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l}), \bar{Q}_\mathrm{L}\}$ of the matrices $\bar{Q}_\mathrm{V}^\top A_\mathrm{V}^\top P_\mathrm{C}$ and $\bar{A}_\mathrm{L} := Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{L}$, respectively, such that

$$\begin{aligned}
V_\mathrm{C}^\top(\mathbf{y}_e, \mathbf{z}_{3e})\hat{C}(\mathbf{y}_e, \mathbf{z}_{3e})\begin{bmatrix} Q_\mathrm{e} & P_\mathrm{e}\end{bmatrix} &= \begin{bmatrix} 0 & I\end{bmatrix}, \\
\bar{V}_\mathrm{L}^\top(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})\hat{L}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})\begin{bmatrix} \bar{Q}_\mathrm{L} & \bar{P}_\mathrm{L}\end{bmatrix} &= \begin{bmatrix} 0 & I\end{bmatrix}.
\end{aligned}$$

The existence of such pairs is guaranteed by Lemma 2.13 (use $N := \hat{C}^{-1}(\mathbf{y}_e, \mathbf{z}_{3e})$, $Q := Q_\mathrm{e}$, $P := P_\mathrm{e}$ for the first pair and and $N := \hat{L}^{-1}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})$, $Q := \bar{Q}_\mathrm{L}$, $P := \bar{P}_\mathrm{L}$ for the second pair). Furthermore, the matrix functions $V_\mathrm{C}^\top$ and $\bar{V}_\mathrm{L}^\top$ are bounded since q_C and ϕ_L are strongly monotone implying $\hat{C}^{-1}$ and $\hat{L}^{-1}$ to be bounded matrix functions [JMT15].

We derive equations of the form (3.9) in four steps, starting with (3.9d) and finishing with (3.9a):

1. Multiplying (3.10a) by $Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top$ and (3.10c) by $\bar{Q}_\mathrm{V}^\top$ from the left yields:

$$\begin{aligned}
\bar{\mathbf{z}}_{3l} &= \bar{M}_{3l}(s_i + s_c(\mathbf{i}_\mathrm{M})), && \text{with } \bar{M}_{3l} := -(Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{L} \bar{P}_\mathrm{L})^{-1} Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top, \\
\mathbf{z}_{3e} &= (M_{3e}^\top M_{3e})^{-1} M_{3e}^\top \bar{Q}_\mathrm{V}^\top s_v, && \text{with } M_{3e} := \bar{Q}_\mathrm{V}^\top A_\mathrm{V}^\top P_\mathrm{C} P_\mathrm{e}.
\end{aligned}$$

   Note, the inverse of $Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{L} \bar{P}_\mathrm{L}$ exists according to Lemma 2.12 (use that $M := Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{L}$ has full row rank due to the absence of IM-cutsets, see Assumption 3.10). Further, M_{3e} has full column rank by the definition of P_e. Introducing $M_3 := \begin{bmatrix} \bar{M}_{3l} & 0 \\ 0 & (M_{3e}^\top M_{3e})^{-1} M_{3e}^\top \bar{Q}_\mathrm{V}^\top \end{bmatrix}$ yields (3.9d).

2. Multiplying (3.10c) by $\bar{P}_\mathrm{V}^\top$ and (3.10a) by $P_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top$ and $P_\mathrm{V}^\top Q_\mathrm{C}^\top$ from the left yields:

$$\begin{aligned}
\mathbf{z}_{2v} &= f_{2v}(\mathbf{y}_e, \mathbf{z}_{3e}, s_v) := (\bar{P}_\mathrm{V}^\top A_\mathrm{V}^\top Q_\mathrm{C} P_\mathrm{V})^{-1} \bar{P}_\mathrm{V}^\top (-A_\mathrm{V}^\top P_\mathrm{C}(Q_\mathrm{e}\mathbf{y}_e + P_\mathrm{e}\mathbf{z}_{3e}) + s_v), \\
\mathbf{z}_{2r} &= f_{2r}(\mathbf{y}, \mathbf{z}_3, s_v, s_i, s_c(\mathbf{i}_\mathrm{M})), \quad \text{with } f_{2r} \text{ satisfying} \\
&\quad h_\mathrm{R}(f_{2r}(\mathbf{y}, \mathbf{z}_3, s_v, s_i, s_c(\mathbf{i}_\mathrm{M})), \mathbf{y}, \mathbf{z}_3, s_v, s_i, s_c(\mathbf{i}_\mathrm{M})) = 0, \\
\bar{\mathbf{z}}_{2v} &= \bar{f}_{2v}(\mathbf{y}, \mathbf{z}_3, s_v, s_i, s_c(\mathbf{i}_\mathrm{M})) \\
&\quad := -(P_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{V} \bar{P}_\mathrm{V})^{-1} P_\mathrm{V}^\top Q_\mathrm{C}^\top [A_\mathrm{R} \hat{g}_\mathrm{R}(f_{2r}(\mathbf{y}, \mathbf{z}_3, s_v, s_i, s_c(\mathbf{i}_\mathrm{M})), f_{2v}(\mathbf{y}_e, \mathbf{z}_{3e}, s_v), \\
&\qquad \mathbf{y}_e, \mathbf{z}_{3e}) + A_\mathrm{L}(\bar{Q}_\mathrm{L}\bar{\mathbf{y}}_l + \bar{P}_\mathrm{L}\bar{\mathbf{z}}_{3l}) + A_\mathrm{V}\bar{Q}_\mathrm{V}\bar{\mathbf{z}}_{1v} + s_i + s_c(\mathbf{i}_\mathrm{M})],
\end{aligned}$$

   where

$$\begin{aligned}
h_\mathrm{R}(\mathbf{z}_{2r}, \mathbf{y}, \mathbf{z}_3, s_v, s_i, s_c(\mathbf{i}_\mathrm{M})) :=& P_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top [A_\mathrm{R} \hat{g}_\mathrm{R}(\mathbf{z}_{2r}, f_{2v}(\mathbf{y}_e, \mathbf{z}_{3e}, s_v), \mathbf{y}_e, \mathbf{z}_{3e}) \\
&+ A_\mathrm{L}(\bar{Q}_\mathrm{L}\bar{\mathbf{y}}_l + \bar{P}_\mathrm{L}\bar{\mathbf{z}}_{3l}) + s_i + s_c(\mathbf{i}_\mathrm{M})].
\end{aligned}$$

   The functions f_{2v} and $\bar{f}_{2v}$ are well-defined since $P_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{V} \bar{P}_\mathrm{V}$ is nonsingular (use again Lemma 2.12 and that $A_\mathrm{V}^\top Q_\mathrm{C} P_\mathrm{V}$ has full column rank by the splitting pair construction). The global Lipschitz continuity of $\hat{g}_\mathrm{R}$ implies f_{2v} and h_R to be globally Lipschitz continuous. Furthermore, h_R is strongly monotone w.r.t. $\mathbf{z}_{2r}$ since $P_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{R}$ is nonsingular (see Lemma 2.12) and g_R is strongly monotone. This ensures the existence of the globally unique function f_{2r} that is also globally Lipschitz continuous [JMT15]. Consequently, also $\bar{f}_{2v}$ is globally Lipschitz continuous. Introducing

$$f_2(\mathbf{y}, \mathbf{z}_3, s_v, s_i, s_c) := \begin{pmatrix} f_{2v}(\mathbf{y}_e, \mathbf{z}_{3e}, s_v) \\ f_{2r}(\mathbf{y}, \mathbf{z}_3, s_v, s_i, s_c) \\ \bar{f}_{2v}(\mathbf{y}, \mathbf{z}_3, s_v, s_i, s_c) \end{pmatrix},$$

   we obtain (3.9c).

3. Multiplying (3.10b) by $\bar{V}_{\mathrm{L}}^{\top}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})$ and (3.10a) by $V_{\mathrm{C}}^{\top}(\mathbf{y}_e, \mathbf{z}_{3e})P_{\mathrm{C}}^{\top}$ from the left yields

$$\mathbf{z}_{1l} = M_{1l}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathbf{z}}_{3l} + f_{1l}(\mathbf{y}, \mathbf{z}_2, \mathbf{z}_3), \tag{3.12a}$$

$$\bar{\mathbf{z}}_{1v} = -\bar{M}_{1v}(\mathbf{y}_e, \mathbf{z}_{3e})\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_{3e} + \bar{f}_{1v}(\mathbf{y}, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c(\mathbf{i}_{\mathrm{M}})) \tag{3.12b}$$

with

$$\begin{aligned}
M_{1l}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l}) :=& (\bar{V}_{\mathrm{L}}^{\top}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})\bar{A}_{\mathrm{L}}^{\top})^{-1},\\
\bar{M}_{1v}(\mathbf{y}_e, \mathbf{z}_{3e}) :=& (V_{\mathrm{C}}^{\top}(\mathbf{y}_e, \mathbf{z}_{3e})P_{\mathrm{C}}^{\top}A_{\mathrm{V}}\bar{Q}_{\mathrm{V}})^{-1},\\
f_{1l}(\mathbf{y}, \mathbf{z}_2, \mathbf{z}_3) :=& -M_{1l}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})\bar{V}_{\mathrm{L}}^{\top}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})\\
&[A_{\mathrm{L}}^{\top}(Q_{\mathrm{C}}(Q_{\mathrm{V}}P_{\mathrm{R}}\mathbf{z}_{2r} + P_{\mathrm{V}}\mathbf{z}_{2v}) + P_{\mathrm{C}}(Q_{\mathrm{e}}\mathbf{y}_e + P_{\mathrm{e}}\mathbf{z}_{3e}))],\\
\bar{f}_{1v}(\mathbf{y}, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c) :=& -\bar{M}_{1v}(\mathbf{y}_e, \mathbf{z}_{3e})V_{\mathrm{C}}^{\top}(\mathbf{y}_e, \mathbf{z}_{3e})P_{\mathrm{C}}^{\top}\\
&[A_{\mathrm{R}}\hat{g}_{\mathrm{R}}(\mathbf{z}_{2r}, \mathbf{z}_{2v}, \mathbf{y}_e, \mathbf{z}_{3e}) + A_{\mathrm{L}}(\bar{Q}_{\mathrm{L}}\bar{\mathbf{y}}_l + \bar{P}_{\mathrm{L}}\bar{\mathbf{z}}_{3l})\\
&+ A_{\mathrm{V}}\bar{P}_{\mathrm{V}}\bar{\mathbf{z}}_{2v} + s_i + s_c(\mathbf{i}_{\mathrm{M}})].
\end{aligned}$$

Note that because of the absence of IM-cutsets, that is by Assumption 3.10, $\bar{A}_{\mathrm{L}}^{\top}$ has full row rank such that from Lemma 2.12 follows the nonsingularity of $\bar{V}_{\mathrm{L}}^{\top}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})\bar{A}_{\mathrm{L}}^{\top}$. Furthermore, $V_{\mathrm{C}}^{\top}(\mathbf{y}_e, \mathbf{z}_{3e})P_{\mathrm{C}}^{\top}A_{\mathrm{V}}\bar{Q}_{\mathrm{V}}$ is nonsingular due to the absence of V-loops. This becomes clear as follows: The absence of V-loops implies A_{V} to have full column rank. By definition of $\bar{Q}_{\mathrm{V}}$, we see that $A_{\mathrm{V}}\bar{Q}_{\mathrm{V}}$ has full column rank as well and that $P_{\mathrm{V}}^{\top}Q_{\mathrm{C}}^{\top}A_{\mathrm{V}}\bar{Q}_{\mathrm{V}} = 0$. By the latter equation and the circumstance that the construction of $Q_{\mathrm{V}}^{\top}$ yields $Q_{\mathrm{V}}^{\top}Q_{\mathrm{C}}^{\top}A_{\mathrm{V}}\bar{Q}_{\mathrm{V}} = 0$, we deduce that $Q_{\mathrm{C}}^{\top}A_{\mathrm{V}}\bar{Q}_{\mathrm{V}} = 0$. Exploiting this and the fact that $\begin{bmatrix}P_{\mathrm{C}}^{\top}\\ Q_{\mathrm{C}}^{\top}\end{bmatrix}A_{\mathrm{V}}\bar{Q}_{\mathrm{V}}$ always has the same column rank as $A_{\mathrm{V}}\bar{Q}_{\mathrm{V}}$, it follows that $P_{\mathrm{C}}^{\top}A_{\mathrm{V}}\bar{Q}_{\mathrm{V}}$ has full column rank. Again, we use Lemma 2.12 for $M := \bar{Q}_{\mathrm{V}}^{\top}A_{\mathrm{V}}^{\top}P_{\mathrm{C}}$ and $P := V_{\mathrm{C}}(\mathbf{y}_e, \mathbf{z}_{3e})$ and are done.

We observe that M_{1l} and $\bar{M}_{1v}$ are bounded matrix functions since $\bar{V}_{\mathrm{L}}^{\top}$ and $V_{\mathrm{C}}^{\top}$ are so. Due to the matrix functions' inverse and several multiplications, only local Lipschitz continuity of f_{1l} and $\bar{f}_{1v}$ can be guaranteed. Introducing

$$\begin{aligned}
M_1(\mathbf{y}, \mathbf{z}_3) &:= \begin{bmatrix}0 & M_{1l}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})\\ -\bar{M}_{1v}(\mathbf{y}_e, \mathbf{z}_{3e}) & 0\end{bmatrix},\\
f_1(\mathbf{y}, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c) &:= \begin{pmatrix}f_{1l}(\mathbf{y}, \mathbf{z}_2, \mathbf{z}_3)\\ \bar{f}_{1v}(\mathbf{y}, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c)\end{pmatrix},
\end{aligned}$$

we obtain (3.9b).

4. Multiplying (3.10b) by $\bar{Q}_{\mathrm{L}}^{\top}$ and (3.10a) by $Q_{\mathrm{e}}^{\top}P_{\mathrm{C}}^{\top}$ from the left and using (3.12) yields

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathbf{y}}_l = \bar{f}_{0l}(\mathbf{y}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) :=\\
&\quad(\bar{Q}_{\mathrm{L}}^{\top}\hat{L}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})\bar{Q}_{\mathrm{L}})^{-1}\bar{Q}_{\mathrm{L}}^{\top}[I - \hat{L}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})\bar{P}_{\mathrm{L}}\bar{V}_{\mathrm{L}}^{\top}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})]\\
&\quad[A_{\mathrm{L}}^{\top}(Q_{\mathrm{C}}(Q_{\mathrm{V}}P_{\mathrm{R}}\mathbf{z}_{2r} + P_{\mathrm{V}}\mathbf{z}_{2v}) + P_{\mathrm{C}}(Q_{\mathrm{e}}\mathbf{y}_e + P_{\mathrm{e}}\mathbf{z}_{3e})) + \bar{A}_{\mathrm{L}}^{\top}\mathbf{z}_{1l}],
\end{aligned}$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}_e &= f_{0e}(\mathbf{y}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c(\mathbf{i}_\mathrm{M})) := \\ &(Q_\mathrm{e}^\top \hat{C}(\mathbf{y}_e, \mathbf{z}_{3e}) Q_\mathrm{e})^{-1} Q_\mathrm{e}^\top P_\mathrm{C}^\top [I - \hat{C}(\mathbf{y}_e, \mathbf{z}_{3e}) P_\mathrm{e} V_\mathrm{C}^\top (\mathbf{y}_e, \mathbf{z}_{3e})] \\ &[A_\mathrm{L}(\bar{Q}_\mathrm{L}\bar{\mathbf{y}}_l + \bar{P}_\mathrm{L}\bar{\mathbf{z}}_{3l}) + A_\mathrm{V}(\bar{Q}_\mathrm{V}\bar{\mathbf{z}}_{1v} + \bar{P}_\mathrm{V}\mathbf{z}_{2v}) \\ &+ A_\mathrm{R}\hat{g}_\mathrm{R}(\mathbf{z}_{2r}, \mathbf{z}_{2v}, \mathbf{y}_e, \mathbf{z}_{3e}) + (s_i + s_c(\mathbf{i}_\mathrm{M}))].\end{aligned}$$

The functions $\bar{f}_{0l}$ and f_{0e} are well-defined since the matrices $\bar{Q}_\mathrm{L}^\top \hat{L}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l})\bar{Q}_\mathrm{L}$ and $Q_\mathrm{e}^\top \hat{C}(\mathbf{y}_e, \mathbf{z}_{3e})Q_\mathrm{e}$ are positive definite. By the same arguments as before, the functions $\bar{f}_{0l}$ and f_{0e} are also local Lipschitz continuous. Introducing

$$f_0(\mathbf{y}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c) := \begin{pmatrix} f_{0e}(\mathbf{y}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c) \\ \bar{f}_{0l}(\mathbf{y}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) \end{pmatrix}$$

yields (3.9a).

Regarding (3.11), the transformation matrix $T = \begin{bmatrix} T_0 & T_1 & T_2 & T_3 \end{bmatrix}$ is given by

$$T_0 = \begin{bmatrix} P_\mathrm{C}Q_\mathrm{e} & 0 \\ 0 & \bar{Q}_\mathrm{L} \\ 0 & 0 \end{bmatrix},\ T_1 = \begin{bmatrix} Q_\mathrm{C}Q_\mathrm{V}Q_\mathrm{R} & 0 \\ 0 & 0 \\ 0 & \bar{Q}_\mathrm{V} \end{bmatrix},\ T_2 = \begin{bmatrix} Q_\mathrm{C}P_\mathrm{V} & Q_\mathrm{C}Q_\mathrm{V}P_\mathrm{R} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{P}_\mathrm{V} \end{bmatrix},\ T_3 = \begin{bmatrix} P_\mathrm{C}P_\mathrm{e} & 0 \\ 0 & \bar{P}_\mathrm{L} \\ 0 & 0 \end{bmatrix}$$

and for the initial condition holds $\mathbf{x}(t_0) = x_0 = T_0 y_0 + T_1\mathbf{z}_1(t_0) + T_2\mathbf{z}_2(t_0) + T_3\mathbf{z}_3(t_0)$. Finally, we can conclude that the system (3.10) is equivalent to the system (3.9) as T is nonsingular by construction. □

For the further analysis it is crucial to introduce some stronger assumptions on the characteristic functions.

Assumption 3.12
The characteristic functions $q_\mathrm{C}, g_\mathrm{R}$ and ϕ_L are additionally chosen so that all functions, that is f_0, f_1 and f_2, of Theorem 3.11 become globally Lipschitz continuously.

Even with Assumption 3.12 to hold, the class of considered circuits is still very big, as most applications consider capacitors and inductors to be linear.

Remark 3.13
In case of linear elements, Assumption 3.12 is automatically fulfilled, as the matrix functions in Theorem 3.11 are no longer depended on arguments.

Given an additional, merely mild, topological restriction, the decoupled expression (3.9) can be simplified even further.

Assumption 3.14
Mock elements do not form a cutset together with inductors and current sources. In other words, there is no LIM^+-cutset.

Assumption 3.14 is a weak topological restriction to the location of mock elements in the circuit and can be understood as a design helper depending on how the overall system is about to be solved, see Chapter 6. Note that a capacitive/resistive path linking the terminals of a mock element is sufficient to meet Assumption 3.14. Moreover, at this point it should be mentioned that LI-cutsets are still allowed.

Corollary 3.15 (Corollary 2 from [ST20])
With Assumption 3.14 to hold, the decoupled system (3.9) simplifies to the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = f_0(\mathbf{y}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c(\mathbf{i}_\mathrm{M})), \tag{3.13a}$$

$$\mathbf{z}_1 = M_1(\mathbf{y}, \mathbf{z}_3)\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3 + f_1(\mathbf{y}, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c(\mathbf{i}_\mathrm{M})), \tag{3.13b}$$

$$\mathbf{z}_2 = f_2(\mathbf{y}, \mathbf{z}_3, s_v, s_i, s_c(\mathbf{i}_\mathrm{M})), \tag{3.13c}$$

$$\mathbf{z}_3 = M_3\begin{pmatrix} s_i \\ s_v \end{pmatrix}. \tag{3.13d}$$

Proof. For the proof we look back to the one of Theorem 3.11. There, we multiplied (3.10a) by $Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top$ and (3.10c) by $\bar{Q}_\mathrm{V}^\top$ from the left yielding:

$$Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top (A_\mathrm{V}\bar{P}_\mathrm{V}\bar{\mathbf{z}}_{2v} + s_i + s_c(\mathbf{i}_\mathrm{M})) = 0. \tag{3.14}$$

Due to Assumption 3.14 we obtain from Corollary 3.5 that $Q_\mathrm{CVR}^\top A_\mathrm{M} = 0$, with Q_CVR being a basis of kernel of $\begin{bmatrix} A_\mathrm{C} & A_\mathrm{V} & A_\mathrm{R} \end{bmatrix}^\top$. Applying Lemma 3.6 twice, we deduce that from $Q_\mathrm{CVR}^\top A_\mathrm{M} = 0$ follows $Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{M} = 0$. Hence, equation (3.14) is equivalent to

$$0 = Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top (A_\mathrm{V}\bar{P}_\mathrm{V}\bar{\mathbf{z}}_{2v} + s_i).$$

Therefore, it is

$$\begin{aligned} \bar{\mathbf{z}}_{3l} &= \bar{M}_{3l}s_i, && \text{with same } \bar{M}_{3l} = -(Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{L}\bar{P}_\mathrm{L})^{-1}Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top, \\ \mathbf{z}_{3e} &= M_{3e}s_v, && \text{with same } M_{3e} = (\bar{Q}_\mathrm{V}^\top A_\mathrm{V}^\top P_\mathrm{C} P_\mathrm{e})^{-1}\bar{Q}_\mathrm{V}^\top, \end{aligned}$$

yielding equation (3.13d). The rest of (3.13) is unchanged. □

Concerning the previous note on resistive/capacitive paths linking each mock element's terminals, we manifest the capacitive variant in the following assumption.

Assumption 3.16
For each mock element, there exists a C-path connecting all of its terminals, i. e. $\operatorname{im} A_\mathrm{M} \subset \operatorname{im} A_\mathrm{C}$.

This topological restriction will be implicitly fulfilled by a certain choice of circuit and EM device coupled variants in Chapter 5. Then, we can even further simplify the equivalent systems (3.13).

Corollary 3.17
With Assumption 3.16 to hold, the decoupled system (3.9) simplifies to the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = f_0(\mathbf{y}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c(\mathbf{i}_\mathrm{M})), \tag{3.15a}$$

$$\mathbf{z}_1 = M_1(\mathbf{y}, \mathbf{z}_3)\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3 + f_1(\mathbf{y}, \mathbf{z}_2, \mathbf{z}_3, s_i, s_c(\mathbf{i}_\mathrm{M})), \tag{3.15b}$$

$$\mathbf{z}_2 = f_2(\mathbf{y}, \mathbf{z}_3, s_v, s_i), \tag{3.15c}$$

$$\mathbf{z}_3 = M_3\begin{pmatrix} s_i \\ s_v \end{pmatrix}. \tag{3.15d}$$

Proof. We notice that the decoupling in Theorem 3.11 is still valid as the requirements are the same. Exploitation of the additional constraint given by Assumption 3.16 simply results in a drop of dependencies, to be seen in (3.15c) and (3.15d). Now, by the definition of Q_C, i. e. $\operatorname{im} Q_\mathrm{C} = \ker A_\mathrm{C}^\top$, it holds $A_\mathrm{C}^\top Q_\mathrm{C} = 0$ which is equivalent to $Q_\mathrm{C}^\top A_\mathrm{C} = 0$. Exploiting the fact that by Assumption 3.16 $\operatorname{im} A_\mathrm{M} \subset \operatorname{im} A_\mathrm{C}$, we obtain as an immediate consequence $Q_\mathrm{C}^\top A_\mathrm{M} = 0$. Since further $s_c(\mathbf{i}_\mathrm{M}) = A_\mathrm{M}\mathbf{i}_\mathrm{M}$, it follows $Q_\mathrm{C}^\top s_c(\mathbf{i}_\mathrm{M}) = 0$.
To explain the losses, we look back to the proof of Theorem 3.11. In the second step, we multiplied (3.10a), that is

$$A_\mathrm{C}\frac{\mathrm{d}}{\mathrm{d}t}q_\mathrm{C}(A_\mathrm{C}^\top\mathbf{e}) + A_\mathrm{R}g_\mathrm{R}(A_\mathrm{R}^\top\mathbf{e}) + A_\mathrm{L}\mathbf{i}_\mathrm{L} + A_\mathrm{V}\mathbf{i}_\mathrm{V} + s_i + s_c(\mathbf{i}_\mathrm{M}) = 0,$$

by $P_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top$ and $P_\mathrm{V}^\top Q_\mathrm{C}^\top$ from the left.

Hence, the mentioned multiplications yield

$$\begin{aligned} P_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top\big[A_\mathrm{R}g_\mathrm{R}(A_\mathrm{R}^\top[Q_\mathrm{C}(Q_\mathrm{V}P_\mathrm{R}\mathbf{z}_{2r} + P_\mathrm{V}\mathbf{z}_{2v}) + P_\mathrm{C}(Q_\mathrm{e}\mathbf{y}_e + P_\mathrm{e}\mathbf{z}_{3e})]) \\ +A_\mathrm{L}[\bar{Q}_\mathrm{L}\bar{\mathbf{y}}_l + \bar{P}_\mathrm{L}\bar{\mathbf{z}}_{3l}] + s_i\big] = 0, \\ P_\mathrm{V}^\top Q_\mathrm{C}^\top\big[A_\mathrm{R}g_\mathrm{R}(A_\mathrm{R}^\top[Q_\mathrm{C}(Q_\mathrm{V}P_\mathrm{R}\mathbf{z}_{2r} + P_\mathrm{V}\mathbf{z}_{2v}) + P_\mathrm{C}(Q_\mathrm{e}\mathbf{y}_e + P_\mathrm{e}\mathbf{z}_{3e})]) \\ +A_\mathrm{L}[\bar{Q}_\mathrm{L}\bar{\mathbf{y}}_l + \bar{P}_\mathrm{L}\bar{\mathbf{z}}_{3l}] + A_\mathrm{V}\bar{P}_\mathrm{V}\bar{\mathbf{z}}_{2v} + s_i\big] = 0. \end{aligned}$$

Analogously to the original decoupling, we equivalently transform these equations into

$$\begin{aligned} \mathbf{z}_{2r} &= f_{2r}(\mathbf{y}, \mathbf{z}_3, s_v, s_i), \quad \text{with } f_{2r} \text{ satisfying } h_\mathrm{R}(f_{2r}(\mathbf{y}, \mathbf{z}_3, s_v, s_i), \mathbf{y}, \mathbf{z}_3, s_v, s_i) = 0, \\ \bar{\mathbf{z}}_{2v} &= \bar{f}_{2v}(\mathbf{y}, \mathbf{z}_3, s_v, s_i) \\ &:= -(P_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{V}\bar{P}_\mathrm{V})^{-1}P_\mathrm{V}^\top Q_\mathrm{C}^\top\big[A_\mathrm{R}\hat{g}_\mathrm{R}(f_{2r}(\mathbf{y}, \mathbf{z}_3, s_v, s_i), f_{2v}(\mathbf{y}_e, \mathbf{z}_{3e}, s_v), \mathbf{y}_e, \mathbf{z}_{3e}) \\ &\quad + A_\mathrm{L}[\bar{Q}_\mathrm{L}\bar{\mathbf{y}}_l + \bar{P}_\mathrm{L}\bar{\mathbf{z}}_{3l}] + A_\mathrm{V}\bar{Q}_\mathrm{V}\bar{\mathbf{z}}_{1v} + s_i\big], \end{aligned}$$

where

$$\begin{aligned} h_\mathrm{R}(\mathbf{z}_{2r}, \mathbf{y}, \mathbf{z}_3, s_v, s_i) :=& P_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top[A_\mathrm{R}\hat{g}_\mathrm{R}(\mathbf{z}_{2r}, f_{2v}(\mathbf{y}_e, \mathbf{z}_{3e}, s_v), \mathbf{y}_e, \mathbf{z}_{3e}) \\ &+ A_\mathrm{L}[\bar{Q}_\mathrm{L}\bar{\mathbf{y}}_l + \bar{P}_\mathrm{L}\bar{\mathbf{z}}_{3l}] + s_i] \\ \hat{g}_\mathrm{R}(\mathbf{z}_{2r}, \mathbf{z}_{2v}, \mathbf{y}_e, \mathbf{z}_{3e}) :=& g_\mathrm{R}(A_\mathrm{R}^\top[Q_\mathrm{C}(Q_\mathrm{V}P_\mathrm{R}\mathbf{z}_{2r} + P_\mathrm{V}\mathbf{z}_{2v}) + P_\mathrm{C}(Q_\mathrm{e}\mathbf{y}_e + P_\mathrm{e}\mathbf{z}_{3e})]) \\ =& g_\mathrm{R}(A_\mathrm{R}^\top\mathbf{e}) \end{aligned}$$

with unchanged

$$\mathbf{z}_{2v} = f_{2v}(\mathbf{y}_e, \mathbf{z}_{3e}, s_v) := (\bar{P}_\mathrm{V}^\top A_\mathrm{V}^\top Q_\mathrm{C} P_\mathrm{V})^{-1} \bar{P}_\mathrm{V}^\top [-A_\mathrm{V}^\top P_\mathrm{C}(Q_\mathrm{e}\mathbf{y}_e + P_\mathrm{e}\mathbf{z}_{3e}) + s_v].$$

As a consequence, we obtain (3.15c) with

$$f_2(\mathbf{y}, \mathbf{z}_3, s_v, s_i) := \begin{pmatrix} f_{2v}(\mathbf{y}_e, \mathbf{z}_{3e}, s_v) \\ f_{2r}(\mathbf{y}, \mathbf{z}_3, s_v, s_i) \\ \bar{f}_{2v}(\mathbf{y}, \mathbf{z}_3, s_v, s_i) \end{pmatrix}.$$

Note that the functions $f_{2r}, \bar{f}_{2v}, h_\mathrm{R}$ and f_2 are still the same but with overloaded dependencies. Next, we look at the first step were we multiplied (3.10a), that is again

$$A_\mathrm{C} \frac{\mathrm{d}}{\mathrm{d}t} q_\mathrm{C}(A_\mathrm{C}^\top \mathbf{e}) + A_\mathrm{R} g_\mathrm{R}(A_\mathrm{R}^\top \mathbf{e}) + A_\mathrm{L}\mathbf{i}_\mathrm{L} + A_\mathrm{V}\mathbf{i}_\mathrm{V} + s_i + s_c(\mathbf{i}_\mathrm{M}) = 0,$$

by $Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top$ from the left. For the same reasons as above it is $Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{M} = 0$ and thus $Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top s_c = 0$ yielding

$$Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top [A_\mathrm{L}\bar{P}_\mathrm{L}\bar{\mathbf{z}}_{3l} + s_i] = 0 \quad \Leftrightarrow \quad \bar{\mathbf{z}}_{3l} = \bar{M}_{3l} s_i$$

with $\bar{M}_{3l} := -(Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{L}\bar{P}_\mathrm{L})^{-1} Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top$. As $\mathbf{z}_{3e}$ stays unchanged, so does M_3 which leads to (3.15d). □

3.4 Conclusions

In this chapter we introduced the MNA modeling approach for electric circuits. Due to its close relation with graph theory we provided new topological interpretation of incidence matrices, developed in [ST20], which are used by the MNA modeling approach for circuits, see Theorem 3.4. Besides the classical lumped elements, such as capacitors, resistors, inductors, voltage sources and current sources, we further introduced a mock element, described by equations (3.5) and (3.6), serving as a placeholder for refined models in Chapter 5. Then, we decoupled the DAE, resulting from the MNA with additional mock elements, in preparation for the convergence analysis in Chapter 6. Here we provided decoupled formulations in Theorem 3.11 and Corollaries 3.15 and 3.17 upon various levels of topological restriction, thereof one which allows the usage of Theorem 3.4.

4 Electromagnetic Devices

As for todays society, it is undeniable that electromagnetic (EM) devices are an essential part of our lives. This is not least due to the fact that we almost always carry such devices around with us, including car keys, credit cards, mobile phones and much more. Right now we are experiencing how almost everything becomes a computer, not to mention the refrigerator that goes on the Internet. All of this realized by the trend that systems on chip become smaller in size whereby the operating frequency increases.

To keep pace with this progress, industry is interested in the development of refined models, see e. g. Figure 4.1, and proper simulation techniques to further enhance these electromagnetic devices while saving expenditures by virtue of laboratory testing, physical prototype creation etc. Lumped models, such as considered in Chapter 3, are idealized and do not take spatially distributed phenomena, e. g. cross-talking and skin effect, into account. This issue can be solved for instance by refined models, see e. g. [Tsu+93; DGL99; DHW04; SDB10; Bau12; Cor+19]. In case of semiconductor materials, for example, the *drift-diffusion model* is frequently used, see e. g. [Sel84; RSM90; Jer12]. Besides this fluid dynamic model, there are also kinetic and quantum models considering even convection or quantum effects, cf. [DMR05; Arg92; Arn+04]. In case of conducting or isolating material, the problem of spatial phenomena neglect can be addressed by using the full set of *Maxwell's equations*, since they describe all macroscopic electromagnetic phenomena, cf. [Str07]. Hereby, Maxwell's equations are completed by *constitutive relations*, linking the field strengths and flux densities by appropriate models derived by empirical observations according to the underlying material.

Depending on the particular application and numerics, various equivalent formulations of Maxwell's equations have been developed in order to solve electromagnetic problems. Besides the classical $\mathbf{E}-\mathbf{H}$ formulation [Jac99], used e. g. in [MNS17], there is for instance the so-called $\mathbf{A}-\boldsymbol{\varphi}$ formulation [Kam90; BP89; Jac99; Str07; WH06; Nol11], the $\mathbf{E}-\boldsymbol{\varphi}$ formulation [Ban10] or $\mathbf{T}-\boldsymbol{\Omega}$ formulation [Car80; Cor20], to name a few. Moreover, Maxwell's equations are often considered in a simplified form to account for the complexity of the underlying problem, see e. g. [HM89; Sch11]. For instance, whereas the full-wave Maxwell's equation approach, or full-Maxwell, considers every field time-varying, the *magneto-quasistatic* (MQS) approach considers electric fluxes as time-invariant and the *electro-quasistatic* (EQS) approach assumes magnetic fluxes to be time-invariant, see e. g. [HM89]. MQS problems, also called *eddy-current problems*, are usually solved when inductive and resistive effects have to be considered whereas capacitive effects are negligible, for instance for machines working at power frequencies [SSH08; BAN00; Bír+90]. EQS is performed when capacitive and resistive effects are of importance and inductive effects are negligible, such as for technical applications which arise from high-voltage technology or microelectronics, see further [Cle+02; Ste+08].

Motivated by the collaboration with *MAGWEL*, a project partner within the EU founded FP7 ICT project *nanoCOPS*, this treatise is dedicated to the magnetic vector potential full-Maxwell approach, the $\mathbf{A} - \boldsymbol{\varphi}$ formulation see [Sch+16; Mat+19]. This approach is widely used and particularly practical when considering semi-conducting media, see e. g. [Che+11; Che+13], and when merging it with the potential framework of electric circuit modeling. However, the $\mathbf{A} - \boldsymbol{\varphi}$ formulation requires additional gauge fixing in order to erase the potentials ambiguity, see for instance [Bau12; CW02b; Jac99; WH06]. Among various available choices of gauge conditions, we chose the Lorenz type gauge condition introduced in [Bau12] as we obtain a system of ordinary differential equations (ODE) after spatial discretization. For further reading on justification of the potential approach in terms of existence and uniqueness, see for instance the works of Alonso, Hiptmair, Nicaise or Valli, e. g. [Nic14; HKO08; AV98]. The $\mathbf{A} - \boldsymbol{\varphi}$ formulation used here is also used for example in [Cor+19; Bau12].

In order to solve electromagnetic problems numerically, the chosen model equations, such as the ones resulting from the $\mathbf{A} - \boldsymbol{\varphi}$ formulation in focus, have to be discretized with respect to both, space and time. Among various possibilities the spatial discretization is done by the *finite integration technique* (FIT), originally introduced by Thomas Weiland [Wei77a]. In a conceptional similar way to the *finite difference time domain* (FDTD) method, a previously developed spatial discretization scheme by Kane Yee [Yee66] which builds upon the differential form of Maxwell's equations, FIT starts with the integral one. Among other discretization techniques such as the *cell method* [Ton01] or the popular *finite element method* [LLC97; Bos98], the FIT with its recent enhancements is preferred by most simulation softwares. One of its notable features is that the deduced *Maxwell's grid equations* (MGEs) exactly solve Maxwell's equations on the computational meshes and that it has good conservation properties, see e. g. [CW01b]. The fields of application for the FIT reach from actuator simulation [FDW07] over radio frequency simulation and microwave simulation [MW07], to mention a few. For a deeper insight to applications we also refer to the FIT review in [Mar02]. Therefore, it is not surprising that the FIT is also the choice of MAGWEL's device-electro-magnetic modeler `DevEM`, a software package whose outstanding feature is the co-simulation of semiconductor devices together with metal interconnect [MAG16], which was a subject of collaboration within the *nanoCOPS* project [Mat+19].
As for the time discretization, typical choices of time integration schemes for the here considered systems are implicit *backward differentiation formula* or *Runge-Kutta methods*, see Chapter 6. At this point it should be mentioned that there are formulations which allow, without too much effort, for the use of explicit time integration schemes, in particular the much favored symplectic *leapfrog integration*, see e. g. [Wei96; Bos99; MNS17]. For an overview of different system formulations we also refer to [Cor+19].

This chapter is constructed as follows. First, we briefly introduce Maxwell's equations in both, differential form and integral form, and discuss their basic properties and consequences. Secondly, we introduce common constitutive relations for insulating, conducting and semi-conducting media. Thirdly, we discuss interface conditions at the surface of different media which in return motivate the deduction of boundary conditions for a finite domain of consideration. Additionally, we provide approaches for excitation and extraction through which the electromagnetic device communicates with the outside world. Thirdly, the just mentioned content is then transferred into the concept of potentials resulting in the $\mathbf{A} - \boldsymbol{\varphi}$ formulation.

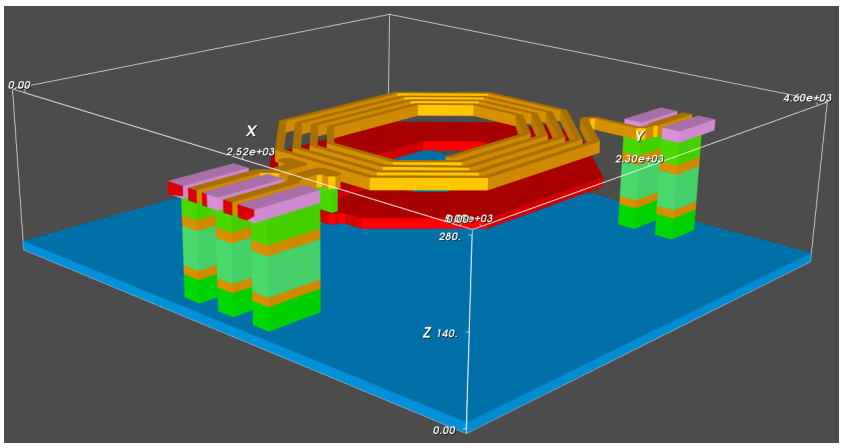


Figure 4.1: Illustration, as taken from `DevEM` [MAG16], shows a simplified prototype balun structure by ACCO Semiconductor, Inc. used in [Sch+16]

Finally, the chapter is concluded by the spatial discretization of the model equations using FIT accompanied by an example of the rectilinear special case without the introduction of phantom cells.

4.1 Maxwell's Equations

With the publication of James Clerk Maxwell's third paper in 1965 [Max65], a new area in the theory of electromagnetism was established. He first described a set of equations unifying the electric and magnetic field theory as classical electromagnetism. These four equations, given in their modern macroscopic formulation in Table 4.1, are nowadays called Maxwell's equations (MEs), see e. g. [Jac99]. The reason why all MEs have individual names is that they are the work of several well-known physicians, but as Maxwell put them together at first and introduced the displacement current, the collection is attributed to him. Gauss's law for magnetism (4.1b) and Maxwell-Faraday's law (4.1c) are referred to as homogeneous MEs and, complementary, Gauss's law (4.1a) and Maxwell-Ampère's law (4.1d) as inhomogeneous MEs.

MEs describe the behavior of four electromagnetic quantities which are expressed by vector-valued functions of space $\Omega \subset \mathbb{R}^3$ and time $\mathcal{I} \subset \mathbb{R}$. These quantities are the *electric* and *magnetic flux densities* $\mathbf{D}, \mathbf{B} : \Omega \times \mathcal{I} \to \mathbb{R}^3$ and the *electric* and *magnetic field strengths* $\mathbf{E}, \mathbf{H} : \Omega \times \mathcal{I} \to \mathbb{R}^3$ depending on the *electric charge density* $\boldsymbol{\rho} : \Omega \times \mathcal{I} \to \mathbb{R}$ and *electric current density* $\mathbf{J} : \Omega \times \mathcal{I} \to \mathbb{R}^3$ which are assumed to be continuous. The quantities' SI units can be looked up in Table 4.3.
Besides the classical form (4.1), an alternative form of MEs is given by the equivalent integral formulation in Table 4.2, which is obtained after applying Gauss's and Stokes' theorem. The integral form is starting point for the *finite integration technique* in the later section. Note that the integral notation is chosen according to convenient physical writing, i. e. the quantity of integral sign indicates the dimension and the ring closeness of the domain. Further, a dot before the integral's delimiter, the differential, indicates that the beforehand field is dot multiplied

name	differential form	
Gauss's law (GL)	$\nabla \cdot \mathbf{D}(r,t) = \boldsymbol{\rho}(r,t)$	(4.1a)
Gauss's law for magnetism (GLM)	$\nabla \cdot \mathbf{B}(r,t) = 0$	(4.1b)
Maxwell-Faraday's law (MF)	$\nabla \times \mathbf{E}(r,t) = -\frac{\partial}{\partial t}\mathbf{B}(r,t)$	(4.1c)
Maxwell-Ampère's law (MA)	$\nabla \times \mathbf{H}(r,t) = \mathbf{J}(r,t) + \frac{\partial}{\partial t}\mathbf{D}(r,t)$	(4.1d)

Table 4.1: Macroscopic Maxwell's equations in modern differential vector-valued formulation using SI unit convention.

with either the according unit normal vector or unit tangential vector of the integral's domain according to the dimension with predefined orientation.

name	integral form	
GL	$\oiint_{\partial V} \mathbf{D}(s,t) \cdot \mathrm{d}s = \iiint_V \boldsymbol{\rho}(r,t)\,\mathrm{d}r$	(4.2a)
GLM	$\oiint_{\partial V} \mathbf{B}(s,t) \cdot \mathrm{d}s = 0$	(4.2b)
MF	$\oint_{\partial S} \mathbf{E}(l,t) \cdot \mathrm{d}l = -\iint_S \frac{\partial}{\partial t}\mathbf{B}(s,t) \cdot \mathrm{d}s$	(4.2c)
MA	$\oint_{\partial S} \mathbf{H}(l,t) \cdot \mathrm{d}l = \iint_S \left(\mathbf{J}(s,t) + \frac{\partial}{\partial t}\mathbf{D}(s,t)\right) \cdot \mathrm{d}s$	(4.2d)

Table 4.2: Macroscopic Maxwell's equations in integral vector-valued and time-domain formulation using SI unit convention.

Current Continuity Equation In the following, we point out some essential properties of MEs. First, the relationship between the current density and the charge density, the so-called *current continuity equation* (CCE) or sometimes called empirical law of the conservation of electric charge

$$\nabla \cdot \mathbf{J}(r,t) + \frac{\partial}{\partial t}\boldsymbol{\rho}(r,t) = 0 \tag{4.3}$$

notation	SI unit	quantity
$\mathbf{A}$	$\frac{\mathrm{Wb}}{\mathrm{m}} = \frac{\mathrm{V\,s}}{\mathrm{m}}$	magnetic vector potential
$\mathbf{B}$	$\mathrm{T} = \frac{\mathrm{V\,s}}{\mathrm{m}^2}$	magnetic flux density
$\mathbf{D}$	$\frac{\mathrm{C}}{\mathrm{m}^2} = \frac{\mathrm{A\,s}}{\mathrm{m}^2}$	electric flux density
$\mathbf{E}$	$\frac{\mathrm{N}}{\mathrm{C}} = \frac{\mathrm{V}}{\mathrm{m}}$	electric field strength
$\mathbf{H}$	$\frac{\mathrm{A}}{\mathrm{m}}$	magnetic field strength
$\mathbf{J}$	$\frac{\mathrm{A}}{\mathrm{m}^2}$	electric current density
$\mathbf{\Pi}$	$\frac{\mathrm{V}}{\mathrm{m}}$	canonical momentum
$\boldsymbol{\rho}$	$\frac{\mathrm{C}}{\mathrm{m}^3} = \frac{\mathrm{A\,s}}{\mathrm{m}^3}$	electric charge density
$\boldsymbol{\varphi}$	V	electric scalar potential

Table 4.3: Electromagnetic quantities and their SI units.

operator	SI unit	name
$\frac{\partial}{\partial t}$	$\frac{1}{\mathrm{s}}$	partial time derivative
∇	$\frac{1}{\mathrm{m}}$	gradient
$\nabla\cdot$	$\frac{1}{\mathrm{m}}$	divergence
$\nabla\times$	$\frac{1}{\mathrm{m}}$	curl

Table 4.4: Differential operators and their SI units.

is satisfied, see e. g. [Jac99; Str07]. This becomes clear by applying the time derivative and divergence on (4.1a) and (4.1d), respectively. The commutation of the operators is admissible given that the fields and their derivatives are continuous. This property states that MEs force the electric charge to be conserved or simply *charge conservation*. Note that applying operators causes a change of units, as to be found in Table 4.4.

Total Current Density With $\mathbf{J}_{\mathrm{tot}} : \Omega \times \mathcal{I} \to \mathbb{R}^3$ we denote the *total current density* flowing through a point in space at a certain time which is defined as the sum of *convection current density* and the electric current density, i. e.

$$\mathbf{J}_{\mathrm{tot}}(r,t) := \frac{\partial}{\partial t}\mathbf{D}(r,t) + \mathbf{J}(r,t). \tag{4.4}$$

Note that $\mathbf{J}$ may also include *external source densities* for instance those arising from coupling other devices or circuits via the boundary.

Lemma 4.1
The total current through a closed surface $\partial\Omega$ equals zero.

Proof. Using the total current density expression (4.4), exploiting MA (4.1d) and Gauss's theorem yields

$$\int_{\partial\Omega} \mathbf{J}_{\mathrm{tot}}(s,t)\cdot \mathrm{d}s = \int_{\partial\Omega} \nabla\times\mathbf{H}(s,t)\cdot \mathrm{d}s = \int_{\Omega} \nabla\cdot\nabla\times\mathbf{H}(r,t)\,\mathrm{d}r = 0$$

due to the identity $\nabla\cdot\nabla\times\equiv 0$. □

Kirchhoff's Circuit Laws Note that MEs are further compatible to Kirchhoff's laws by meaning that they can be deduced by stationary MEs, see e. g. [Nag75]. This property is quite important with regard to the coupling of EM devices into circuits in Chapter 5.

4.2 Constitutive Relations

As it was shown in the previous section, Maxwell's equations (4.1) consist of two curl and two divergence equations for the unknown quantities $\mathbf{E},\mathbf{B},\mathbf{D},\mathbf{H},\mathbf{J}$ and $\boldsymbol{\rho}$. Even though it is well-known that in this formulation are hidden redundancies, (4.1) is not yet determinate - and neither is (4.2) [Str07]. On analogy with the constitutive element equations of electric circuits in Chapter 3, we link the strengths, fluxes and sources by additional so-called *constitutive relations* which then complete Maxwell's equations. This can be realized, for instance, by deducing $\mathbf{D}$, $\mathbf{H}$ and $\mathbf{J}$ from $\mathbf{E}$ and $\mathbf{B}$ based on empirical observations when materials come into play whereby the sources, i. e. $\mathbf{J}$ and $\boldsymbol{\rho}$, always have to comply the continuity equation (4.3). Depending on the field of application, the constitutive relations can be very complex, e. g. nonlinear, nonlocal (in space and time) or frequency and wave vector dependent, cf. [Jac99], that is

$$\mathbf{D}(r,t) = \mathscr{D}[\mathbf{E},\mathbf{B},r,t,\omega_r,\omega_t\ldots](r,t), \tag{4.5a}$$
$$\mathbf{H}(r,t) = \mathscr{H}[\mathbf{E},\mathbf{B},r,t,\omega_r,\omega_t,\ldots](r,t), \tag{4.5b}$$
$$\mathbf{J}(r,t) = \mathscr{J}[\mathbf{E},\mathbf{B},r,t,\omega_r,\omega_t,\ldots](r,t), \tag{4.5c}$$

for some operators $\mathscr{D}$, $\mathscr{H}$ and $\mathscr{J}$. The latter equation (4.5c) is referred to as generalized *Ohm's law*. In this treatise we neglect most of the sophisticated dependencies, such as past (hysteresis), dispersion (by virtue of the presence of $\frac{\partial}{\partial t}$ and ∇), temperature etc. but refer to [WL03] and [ML09] for further readings. Usually, assuming the media to be nonlocal is very important, since they have to follow the *principle of causality*.

Remark 4.2 (principle of causality)
The formulation of constitutive equations for realistic materials must follow the principle of causality, by means no material medium can have an instantaneous response and can not respond before it is stimulated. This would imply that only the past should be involved, since cause and effect can not happen simultaneously.

Insulating and Conducting Media To clarify the level of modeling that is covered in the forthcoming part of the treatise, for insulating and conducting media we consider the constitutive relations

$$\mathbf{D}(r,t) = \varepsilon(r)\mathbf{E}(r,t), \qquad \mathbf{H}(r,t) = \nu(r)\mathbf{B}(r,t), \qquad \mathbf{J}(r,t) = \sigma(r)\mathbf{E}(r,t) \tag{4.6}$$

with *permittivity* $\varepsilon : \Omega \to \mathbb{R}^{3\times3}$, *reluctivity* $\nu : \Omega \to \mathbb{R}^{3\times3}$ and *conductivity* $\sigma : \Omega \to \mathbb{R}^{3\times3}$ being rank-2-tensors, see e. g. [Jac99]. It is very common to write $\mu^{-1} = \nu$ whereby μ is the *permeability*, in this case also a rank-2 tensor. The relations (4.6) cover for instance linear, inhomogeneous and isotropic materials. For more complex media they need to be extended, see for instance [WL03]. As for the upcoming analysis it is imperative to make the following assumption.

Assumption 4.3 (material properties)
The permittivity ε and reluctivity ν are symmetric positive definite and the conductivity σ is symmetric positive semi-definite.

In insulating media, such as for isolators, the conductivity is set to zero. For the sake of readability, the spatial and time arguments will be dropped from now on if not explicitly needed.

Semi-Conducting Media At no time, the global validity of MEs is affected, the only things that change are the source models for $\mathbf{J}$ and $\boldsymbol{\rho}$ which now incorporate doping profiles manipulating the charge densities. In semi-conducting media, the constitutive relation for the sources $\mathbf{J}$ and $\boldsymbol{\rho}$ can be realized by a drift-diffusion model if a few assumptions about the media are made, see further [Sel84; Mar86]. For that, the charge density $\boldsymbol{\rho}$ is substituted by the elementary charge q times the sum of positively charged hole concentration $\mathbf{p} : \Omega \times \mathcal{I} \to \mathbb{R}$, negative charged electron concentration $\mathbf{n} : \Omega \times \mathcal{I} \to \mathbb{R}$ and the typically time invariant donator and exceptor concentrations N_D^+ and $-N_A^- : \Omega \to \mathbb{R}$, i. e.

$$\boldsymbol{\rho} = q(\mathbf{p} - \mathbf{n} + N_D^+ - N_A^-). \tag{4.7}$$

Donator and exceptor concentrations can be grouped to $C := N_D^+ - N_A^-$ [cm^{-3}], the so-called *doping profile*. The current density $\mathbf{J}$ is then interpreted as the sum of current densities caused by holes and electrons, respectively named $\mathbf{J_p}$ and $\mathbf{J_n}$, i. e.

$$\mathbf{J} = \mathbf{J_p} + \mathbf{J_n}.$$

As shown earlier, MEs are compatible to the current continuity equation. Therefore, we supplement the just introduced models with the current continuity equation (4.3), i. e.

$$\nabla \cdot (\mathbf{J_p} + \mathbf{J_n}) + q\frac{\partial}{\partial t}(\mathbf{p} - \mathbf{n}) = 0, \tag{4.8}$$

We can separate the equation into two continuity equations by

$$\nabla \cdot \mathbf{J_p} + q\frac{\partial}{\partial t}\mathbf{p} = -qR(\mathbf{p},\mathbf{n}), \tag{4.9a}$$

$$\nabla \cdot \mathbf{J_n} - q\frac{\partial}{\partial t}\mathbf{n} = qR(\mathbf{p},\mathbf{n}) \tag{4.9b}$$

where R is a suitable model function of $\mathbf{p}$ and $\mathbf{n}$ describing the balance of net generation and recombination of electrons and holes. We assume that this model function is known a priory which further should be handled carefully, cf. [Sel84] or [Lun00]. Note that with (4.8) to hold (4.9) is satisfied. Without loss of generality, the current densities of charged particles, regarding the negative sign convention, can be written as

$$\mathbf{J_p} = q\mathbf{p}v_\mathrm{p}, \qquad \mathbf{J_n} = -q\mathbf{n}v_\mathrm{n}$$

where v_p and v_n are average drift velocity, see e. g. [Sel84]. Depending on the choice of average velocities, we obtain different models. We conclude with a classical model for these current relations

$$\mathbf{J_p} \cong q\mu_\mathrm{p}\left(\mathbf{pE} - \frac{k\bar{T}}{q}\nabla\mathbf{p}\right), \tag{4.10a}$$

$$\mathbf{J_n} \cong q\mu_\mathrm{n}\left(\mathbf{nE} + \frac{k\bar{T}}{q}\nabla\mathbf{n}\right) \tag{4.10b}$$

which is widely-accepted and consists of a drift and diffusion component, see also [MSM01]. The parameters μ_p and μ_n are representing the effective carrier mobilities, depending on the underlying material, chosen by empirical observation. Again, k is the Boltzmann constant and $\bar{T}$ the nominal temperature parameter as already introduced for the diode in Chapter 3. Justifying this modeling approach (4.7)-(4.10) requires a bunch of assumptions, see e. g. [Sel84], which can be summarized as in Assumption 4.4.

Assumption 4.4 (semiconductor modeling)

1. *All scattering processes are assumed to be elastic.*
2. *Spatial variation of the collision time and the band structure are neglected.*
3. *Effects of degeneracy are neglected in the approximation for the scattering integral.*
4. *Spatial variation of the external force is ignored.*
5. *Influence of Lorentz force is ignored.*
6. *Time and spatial variation of carrier temperature is neglected and assumed to be equal to the lattice temperature.*
7. *Parabolic energy bands are assumed.*
8. *Zero order term of the series expansion of* $\mathbf{J_p}$ *and* $\mathbf{J_n}$ *into powers of the collision time only taken into account.*
9. *Semiconductor is assumed to be infinitely large.*

Extra assumptions for the derivation of the current densities grad-form in semiconductors:

1. *Higher order derivatives of the quasi-Fermi potentials are neglected.*
2. *Dependence of the distribution function upon the gradient of the quasi-Fermi potential is linearized.*

4.3 Interface and Boundary Conditions

In this section we cover what happens with the electromagnetic fields at interfaces between different media, referred to as *interface conditions*, and at the boundary of a finite domain of consideration, referred to as *boundary conditions*.
Whereas inside a medium the material constitutive relations are usually smooth or even constant, we encounter sharp changes across the surface which separates different media. Hence, in a macroscopic point of view, one may usually consider them as discontinuities.
As one of the four conventionally accepted fundamental interactions, electromagnetism is infinitely ranged alongside with gravity. Since, we can only consider finite spaces in terms of simulations, one defines a bounded domain for the electromagnetic device and provides *artificial boundary conditions* which arise from assuming idealistic material surrounding the device. These boundary conditions can somehow be interpreted as a special case of interface conditions taking into account asymptotic behavior of the electromagnetic quantities on one side, see (4.12). In addition to that, an EM device usually communicates with the outside world via contacts at its boundary. This communication can be realized by incorporating additional excitations or extractions at these so-called *physical boundaries*.
For clarity, there is given a schematic overview of the interfaces and boundaries in Figure 4.5. Note that for the sake of human readability, the spatial and time arguments are omitted.

4.3.1 Media Interface Conditions

Starting from two neighboring media originated in the disjoint regions Ω_1 and Ω_2, we denote with $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$ the surface that is shared by both regional boundaries. Further, it is n_Γ the interfaces unit normal vector being exterior to Ω_1 without loss of generality.
Let V be an arbitrary finite volume with piecewise smooth surface S and outer unit normal n, straddling the interface region qualitatively fitting to the box given in Figure 4.2. We now

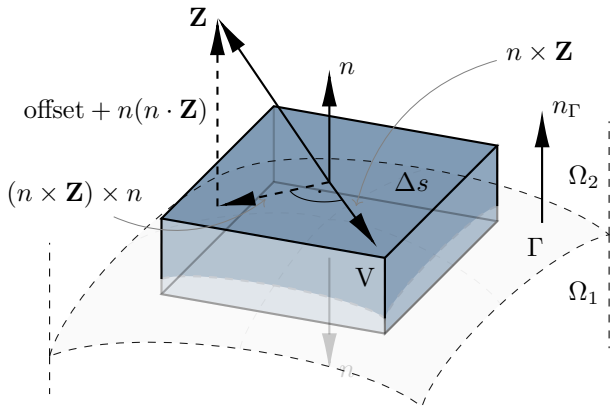


Figure 4.2: Skizzes of volume V straddling the interface Γ between different media originated in Ω_1 and Ω_2 and the dissection of an arbitrary field $\vec{Z}$.

shallow the volume so that the contribution of its side surfaces can be neglected and only

top and bottom of the surface do matter. Assume that for some area Δs top and bottom are now parallel and tangential to the interface surface. We notice that the volume's top and bottom exterior unit normals equal the surface's unit normal and its inverse, respectively, i. e. $n|_{\Omega_2} = n_\Gamma$ and $n|_{\Omega_1} = -n_\Gamma$. Hence, we obtain respectively from GL (4.2a) and GLM (4.2b) for this area the following left-hand sides

$$\int_S \mathbf{D} \cdot n \, \mathrm{d}s = (\mathbf{D}|_{\Omega_2} - \mathbf{D}|_{\Omega_1}) \cdot n_\Gamma \Delta s \quad \text{and} \quad \int_S \mathbf{B} \cdot n \, \mathrm{d}s = (\mathbf{B}|_{\Omega_2} - \mathbf{B}|_{\Omega_1}) \cdot n_\Gamma \Delta s.$$

For the right-hand sides we obtain for Gauss's law for magnetism trivially zero and for Gauss law

$$\int_V \boldsymbol{\rho} \, \mathrm{d}r = \boldsymbol{\rho}_s \Delta s$$

with $\boldsymbol{\rho}_s$ being the *idealized surface charge density*. More details can be found in [Jac99] or [Str07]. Analogously, let C be a arbitrary closed and piecewise smooth contour, with unit tangential field τ, spanning a regular surface S with unit normal field n. Applying the same procedure on the contour C, which goes through both media as shown in Figure 4.3, yields the relation $\tau|_{\Omega_2} = -\tau|_{\Omega_1} =: \tau_\Gamma$ in the limit of flattening. Hence, in case of an infinitesimally top

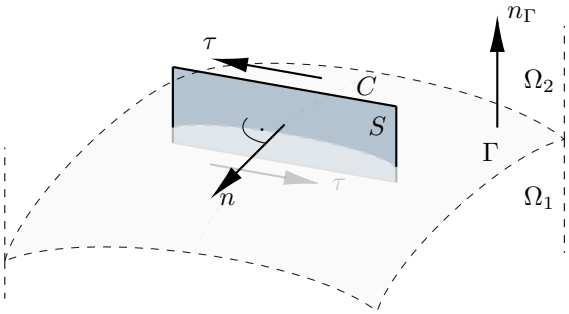


Figure 4.3: Skizzes of contour C and its spanned surface S straddling the interface Γ between different media originated in Ω_1 and Ω_2. The tangent field τ is orthogonal to the unit normal field n in each point.

and bot line segment, straight and parallel to the surface of length Δl, where the end segments' contribution vanishes, we obtain for the left integral expressions of MF (4.2c) and MA (4.2d)

$$\int_C \mathbf{E} \cdot \tau \, \mathrm{d}l = (\mathbf{E}|_{\Omega_2} - \mathbf{E}|_{\Omega_1}) \cdot \tau_\Gamma \Delta l, \quad \text{and} \quad \int_C \mathbf{H} \cdot \tau \, \mathrm{d}l = (\mathbf{H}|_{\Omega_2} - \mathbf{H}|_{\Omega_1}) \cdot \tau_\Gamma \Delta l.$$

The vector field derivative on the right-hand side vanishes with Δl, if the fields and their derivatives are bounded. Thus, the only term remaining is current density term of Maxwell-Ampeères law so that we can write

$$\int_S \mathbf{J} \cdot n \, \mathrm{d}s = \mathbf{J}_s \cdot n_\Gamma \Delta l,$$

for some so-called *idealized surface density* $\mathbf{J}_s$. In the limiting case we can also express $\tau_\Gamma = n_\Gamma \times n$. Thus, we collect the boundary conditions, expressed in terms of the interfaces unit normal n_Γ, for each point of the interface as follows:

$$(\mathbf{D}|_{\Omega_2} - \mathbf{D}|_{\Omega_1}) \cdot n_\Gamma = \boldsymbol{\rho}_s, \tag{4.11a}$$
$$(\mathbf{B}|_{\Omega_2} - \mathbf{B}|_{\Omega_1}) \cdot n_\Gamma = 0, \tag{4.11b}$$
$$(\mathbf{E}|_{\Omega_2} - \mathbf{E}|_{\Omega_1}) \times n_\Gamma = 0, \tag{4.11c}$$
$$(\mathbf{H}|_{\Omega_2} - \mathbf{H}|_{\Omega_1}) \times n_\Gamma = \mathbf{J}_s, \tag{4.11d}$$

These conditions can be collected in a graphic, see Figure 4.4.

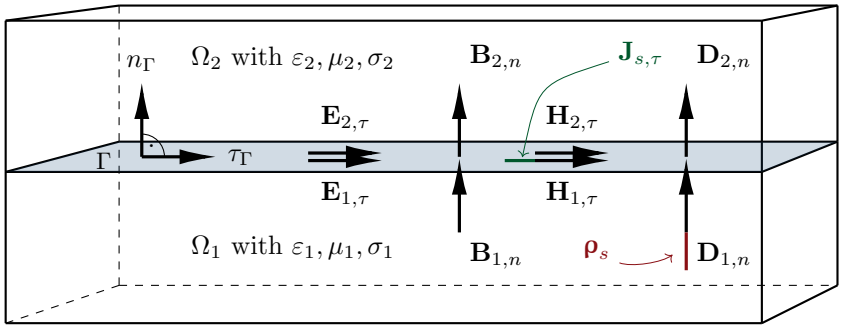


Figure 4.4: Interface condition visualization. Whereas the numeric subscripts 1 and 2 indicate quantities reduced to the regions Ω_1 and Ω_2, n and τ refer to the normal and the tangential component in direction of n_Γ and τ_Γ, respectively.

The above interface conditions are media independent, thus they hold also for semi-conducting media. In the latter case, there are just the exceptions that additional conditions have to be found for the electron and hole concentrations $\mathbf{n}$ and $\mathbf{p}$.

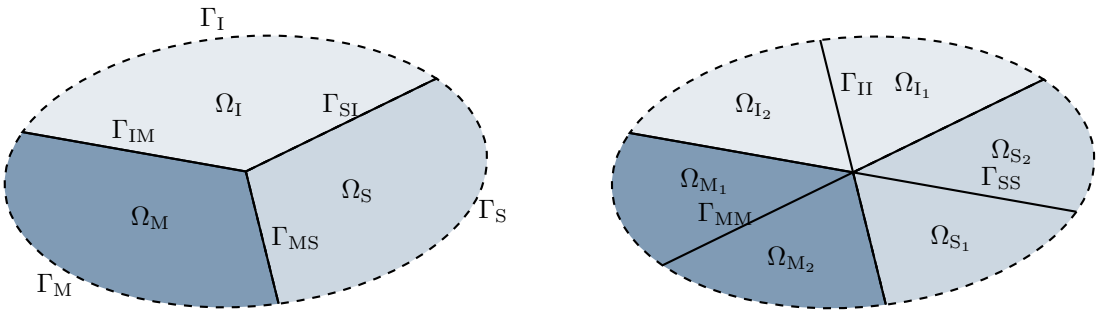


Figure 4.5: Possible boundaries and interfaces for simulation domain. The domain is split in three groups of media: Ω_M for conducting, Ω_S for semi-conducting and Ω_I for insulating ones. Dashed lines, i. e. Γ_M, Γ_S and Γ_I are boundaries and solid lines represent interfaces.

Again, following the idea of distinguishing media into three classes of media, i. e. conducting, insulating and semi-conducting, we obtain the following possible combinations of interfaces, see Figure 4.5,

- conductor-conductor
- conductor-insulator
- insulator-insulator
- semiconductor-conductor
- semiconductor-insulator
- semiconductor-semiconductor

The first three interface types, conductor-conductor, conductor-insulator and insulator-insulator are perfectly covered by (4.11). Sometimes the electric permittivity for the insulator is assumed to vanish for simplicity, even though it is lower bounded by the electric permeability in vacuum $\varepsilon_0 \approx 8.8541878128 \cdot 10^{-12}$.
Interfaces with at least one medium being semi-conducting are a little bit more complex, since there quantities change, i. e. electron and hole concentrations, $\mathbf{n}$ and $\mathbf{p}$, are consequences of semiconductor modeling only. Following [Sel84, p. 133] we have for instance that the current density components perpendicular to the interface with an insulator have to equal the a priori given surface recombination rate R^{surf}, i. e.

$$\begin{aligned} \mathbf{J_n} \cdot n_\Gamma &= -qR^{\text{surf}}, \\ \mathbf{J_p} \cdot n_\Gamma &= qR^{\text{surf}}. \end{aligned}$$

Usually, this interface is assumed to be ideal, by meaning that the surface recombination rate R^{surf} is zero.

4.3.2 Spatial Boundary Conditions

Theoretically, the domain of consideration Ω for MEs (4.1) is unbounded, i. e. $\Omega = \mathbb{R}^3$, since the electromagnetic influence is infinity. In practice, for example when numerically approximating, it is typically restricted to a finite space.

Assumption 4.5
Let $\Omega \subset \mathbb{R}^3$ be bounded and simply connected with $\Gamma = \partial\Omega$ being its boundary.

In order to obtain a unique solution there must be at least one Dirichlet boundary part among Γ, see for instance [Fri08; QV08]. The boundaries are generally split up as $\partial\Omega = \partial\Omega^{\text{phys}} \cup \partial\Omega^{\text{art}}$ into the so-called *physical boundaries* and *artificial boundaries*. The former is represented by a finite number of contacts, setting the device into relation with some environment, for example electronic circuits, discussed in Chapter 3. Artificial boundaries are used to simply cut of the device's domain where the electromagnetic influence becomes negligibly small and thus is not of interest for computation. Concerning semiconductors, they usually also describe parts where the device has been separated from an adjacent one, e. g. in integrated circuits.

Artificial Boundaries In literature there can be found a bunch of conditions describing artificial boundaries which are discussed for instance in [AR03; Ben06; AV10; Sch11; Cor+19]. Typically, the modeled device is assumed to be surrounded by materials with idealized electric or magnetic property, or combinations thereof. For these common boundary types, let $\Gamma = \partial\Omega^{\text{art}}$ decompose in Γ_{E} and Γ_{H}, such that

$$\Gamma = \bar{\Gamma}_E \cup \bar{\Gamma}_H, \qquad \Gamma_{\text{E}} \cap \Gamma_{\text{H}} = \emptyset.$$

Let n be the outer unit normal on Γ then these conditions, deduced by (4.11), read

$$\mathbf{E} \times n = \mathbf{E}_0 \times n \quad \text{on } \Gamma_{\text{E}} \times \mathcal{I}, \tag{4.12a}$$
$$\mathbf{H} \times n = \mathbf{H}_0 \times n \quad \text{on } \Gamma_{\text{H}} \times \mathcal{I} \tag{4.12b}$$

where $\mathbf{E}_0$ and $\mathbf{H}_0$ describe the fields trends and assumed to be known. For idealized materials they become zero, yielding the following boundary conditions

- *Electric boundary condition*, assuming a perfectly electrically conducting (PEC) medium at boundary ($\sigma \to \infty$), corresponding to a vanishing tangential component of the electric field at the boundary, i. e. $\mathbf{E}_0 \times n = 0$ in (4.12a) yielding

$$\mathbf{E} \times n = 0 \quad \text{on } \Gamma_{\text{E}} \times \mathcal{I}, \tag{4.13a}$$
$$\mathbf{B} \cdot n = 0 \quad \text{on } \Gamma_{\text{E}} \times \mathcal{I}. \tag{4.13b}$$

  Motivated by the first equation, the latter condition is obtained by Lemma A.1.

- *Magnetic boundary condition*, assuming a perfectly magnetically permeable, or also called perfectly magnetically conducting, medium at the boundary ($\mu \to \infty$ or $\nu \to 0$), corresponding to a vanishing tangential component of the magnetic field at the boundary, i. e. $\mathbf{H}_0 \times n = 0$ in (4.12b) yielding

$$\mathbf{H} \times n = 0 \quad \text{on } \Gamma_{\text{H}} \times \mathcal{I},$$
$$\mathbf{D} \cdot n = 0 \quad \text{on } \Gamma_{\text{H}} \times \mathcal{I}.$$

  The latter condition is obtained analogously to Lemma A.1 by making use of Maxwell-Ampére instead and assuming $\mathbf{J} \cdot n = 0$.

Other possibilities are for example the so-called *open boundary*, *periodic boundary* or *anti-periodic boundary* approaches, read further [Cle98; AV10; Sch11]. An advantage of the electric boundary condition over the magnetic boundary condition is that it yields a *Dirichlet boundary condition* for the $\mathbf{A} - \boldsymbol{\varphi}$ formulation introduced in Section 4.4. The magnetic boundary condition however yields a *Neumann boundary condition* which is why we focus ourselves on the PEC case. Note that commonly imposed boundary conditions produce spurious reflections from the surrounding confining region. In order to avoid such reflections, researchers have spend years on developing suitable conditions, especially with regard to numerical approximation, such as nonreflecting boundary or transparent boundary conditions, see further [Giv91; GK96; CN08].

For semiconductors, the device is assumed to be self-contained, that is the outgoing components of the electric field strength and the current density vanish, see e. g. [Sel84; Mar86; Bau+13], which corresponds to the Neumann boundary conditions

$$\begin{aligned} \nabla \mathbf{n} \cdot n &= 0 \quad \text{on } \Gamma_S \times \mathcal{I}, \\ \nabla \mathbf{p} \cdot n &= 0 \quad \text{on } \Gamma_S \times \mathcal{I} \end{aligned}$$

whereby Γ_S may share some boundary part with Γ_E or Γ_H. For the current densities we have

$$\begin{aligned} \mathbf{J_n} \cdot n &= 0 \quad \text{on } \Gamma_S \times \mathcal{I}, \\ \mathbf{J_p} \cdot n &= 0 \quad \text{on } \Gamma_S \times \mathcal{I} \end{aligned}$$

and hence $\mathbf{J} \cdot n = \mathbf{J_n} \cdot n + \mathbf{J_p} \cdot n = 0$.

Physical Boundaries When simulation electromagnetic devices in a bounded domain, they usually communicate with the external environment by some kind of excitations or extractions, which can be understood as some kind of input/output system. Contacts can be voltage, current or mixed current-voltage driven and are usually assumed to be perfectly conducting which is why they comply with the artificial electric boundary condition and an additional excitation or extraction.

Let $\Gamma \subset \partial\Omega^{\text{phys}}$ be a contact. Expressed in potentials, see (4.16) and (4.17), the electric boundary potential can be excited by an applied potential $\varphi_\Gamma : \mathcal{I} \to \mathbb{R}$ as follows

$$\boldsymbol{\varphi} = \varphi_\Gamma(t) \quad \text{on } \Gamma \times \mathcal{I}. \tag{4.14}$$

The total current through that same surface is, according to (4.4), then given by the function $i_\Gamma : \mathcal{I} \to \mathbb{R}$ complying

$$i_\Gamma(t) = \int_\Gamma \mathbf{J}_{\text{tot}} \cdot \mathrm{d}s = \int_\Gamma \left[\frac{\partial}{\partial t}\mathbf{D} + \mathbf{J}\right] \cdot \mathrm{d}s \quad \text{on } \mathcal{I}. \tag{4.15}$$

For semiconductors there are various types of contact models, e. g. Ohmic contacts, Shottky contacts or Polysilicon contacts. Due to a small Schottky barrier hight, an Ohmic contact has negligible contact resistance, and therefore is preferred in most applications. Criteria for a good Ohmic contact can be found for instance in [Pio93]. These boundary conditions formally read

$$g(\boldsymbol{\varphi}, \frac{\partial}{\partial t}\boldsymbol{\varphi}, i_\Gamma, \frac{\partial}{\partial t} i_\Gamma) = 0 \quad \text{on } \Gamma \times \mathcal{I}.$$

For some given function $g : \mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ and a simplified current $i_\Gamma(t) = \int_\Gamma \left[\mathbf{J_p} + \mathbf{J_n}\right] \cdot \mathrm{d}s$. In a special case of a purely voltage driven contact, we set the potential $\boldsymbol{\varphi}$, similar to (4.14) as follows:

$$\boldsymbol{\varphi} = \varphi_{\text{bi}} + \phi_\Gamma(t) \quad \text{on } \Gamma \times \mathcal{I}$$

where φ_{bi} is the built-in potential. In case of purely current driven contacts we can proceed as in (4.15) whereby $\mathbf{J} = \mathbf{J_p} + \mathbf{J_n}$. Additional conditions for the carrier concentrations can, for given n_i, be expressed by the Dirichlet boundary conditions, see for instance [Sel84],

$$\begin{aligned} \mathbf{p} &= \frac{1}{2}\left(\sqrt{C^2 + 4n_i^2} - C\right) \quad \text{on } \Gamma \times \mathcal{I}, \\ \mathbf{n} &= \frac{1}{2}\left(\sqrt{C^2 + 4n_i^2} + C\right) \quad \text{on } \Gamma \times \mathcal{I}. \end{aligned}$$

4.4 A − φ Formulation

Instead of solving MEs in their classical form (4.1), a system consisting of four first order partial differential equations (PDEs), one can also first transform them into another equivalent formulation. Other formulations are for instance the $\mathbf{A} - \boldsymbol{\varphi}$ formulation [Kam90; BP89; Jac99; Str07; WH06; Nol11], with the $\mathbf{A}^*$ formulation as a special case [ET88], the $\mathbf{E} - \boldsymbol{\varphi}$ formulation [Ban10] or the $\mathbf{T} - \boldsymbol{\Omega}$ formulation [Car80], depending on the application and what is more efficient to solve under the given circumstances. Here we focus on the so-called $\mathbf{A} - \boldsymbol{\varphi}$ formulation obtained after rewriting MEs in terms of potentials which are the *magnetic vector potential* $\mathbf{A}$ and *electric scalar potential* $\boldsymbol{\varphi}$, leading to two second order PDEs.
With the introduction of these potentials, the homogeneous MEs, namely Gauss's law for magnetism (4.1b) and Maxwell-Faraday's law (4.1c), are fulfilled implicitly and it remains to solve Gauss's law (4.1a) and Maxwell-Ampère's law (4.1d). This approach, on the other hand, involves ambiguity of the potentials which makes additional gauge fixing necessarily, see for instance [Jac99; WH06; CW02b; Bau12].

From the Gauss's law for magnetism (4.1b) we know that $\mathbf{B}$ is a solenoidal vector field. Thus, using Helmholtz decomposition, the magnetic field can be expressed as the curl of some other vector field which will be the magnetic vector potential $\mathbf{A} : \Omega \times \mathcal{I} \to \mathbb{R}^3$ fulfilling

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{4.16}$$

Substituting (4.16) in Maxwell-Faraday's law yields

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B} = -\frac{\partial}{\partial t}\nabla \times \mathbf{A} = -\nabla \times \frac{\partial}{\partial t}\mathbf{A} \quad \Rightarrow \quad \nabla \times \left(\mathbf{E} + \frac{\partial}{\partial t}\mathbf{A}\right) = 0.$$

Thus, we can express this quantity using the gradient of another scalar field $\boldsymbol{\varphi} : \Omega \times \mathcal{I} \to \mathbb{R}$, the so-called electric scalar potential, satisfying

$$\mathbf{E} = -\nabla\boldsymbol{\varphi} - \frac{\partial}{\partial t}\mathbf{A}. \tag{4.17}$$

With these auxiliary equations for the potentials, (4.16) and (4.17), the homogeneous MEs are trivially fulfilled by construction, i. e. taking the gradient of (4.16) and (4.17) implies (4.1b) and (4.1c), respectively. Thus, we must use the inhomogeneous MEs in order to determine the

potentials. Exploiting further the constitutive relations (4.6), Maxwell's equations turn to

$$-\nabla \cdot \left[\varepsilon\left(\nabla \boldsymbol{\varphi} + \frac{\partial}{\partial t}\mathbf{A}\right)\right] = \boldsymbol{\rho}, \tag{4.18a}$$

$$\nabla \times (\nu \nabla \times \mathbf{A}) + \frac{\partial}{\partial t}\left[\varepsilon\left(\nabla \boldsymbol{\varphi} + \frac{\partial}{\partial t}\mathbf{A}\right)\right] = \mathbf{J} \tag{4.18b}$$

where the sources remain to be chosen in accordance with the underlying material and are related to each other by the current continuity equation. This equivalent system of equations (4.18) is referred to as the $\mathbf{A} - \boldsymbol{\varphi}$ *formulation* of MEs, see further [Bos98]. Figure 4.6 shows the so-called *Maxwell's house* or *Tonti's diagram*, see e. g. [Bos91; Ton75; Ton95; Cle05; Des81], which provides an overview of the electromagnetic quantities, how they are connected and translate into each other.

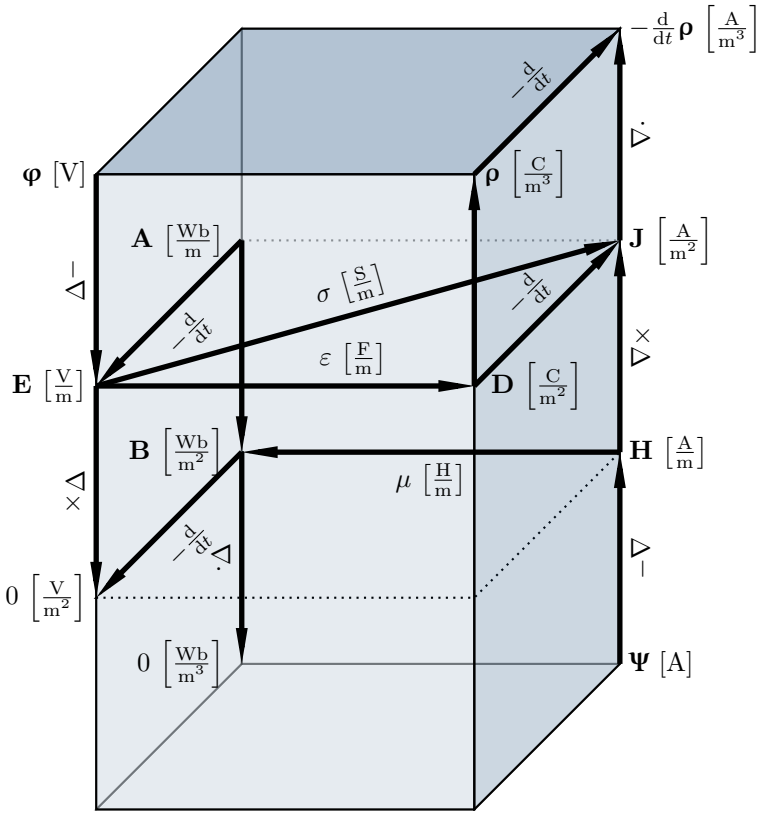


Figure 4.6: Maxwell's house or Tonti's diagram for continuous electromagnetic quantities.

Quasi-Canonical Momentum In order to avoid second order derivatives, we additionally introduce the so-called *quasi-canonical momentum* $\boldsymbol{\Pi} : \Omega \times \mathcal{I} \to \mathbb{R}^3$, as used for instance in [Sch+16; Che+13], satisfying

$$\boldsymbol{\Pi} = \frac{\partial}{\partial t}\mathbf{A}. \tag{4.19}$$

4.4.1 Boundary Conditions

A possible choice for the $\mathbf{A} - \boldsymbol{\varphi}$ formulation's (Dirichlet) boundary conditions can be deduced from assuming PEC boundaries, see e. g. [HA01], as follows.

Assumption 4.6 (boundary conditions for $\mathbf{A} - \boldsymbol{\varphi}$ formulation)
We assume PEC medium at the boundary so that

$$\begin{aligned} \mathbf{A} \times n &= 0 \quad \text{on } \partial\Omega \times \mathcal{I}, \\ \nabla\boldsymbol{\varphi} \times n &= 0 \quad \text{on } \partial\Omega \times \mathcal{I}. \end{aligned}$$

4.4.2 Gauge Fixing

With this widely used potential approach, hidden redundancies in Maxwell's equations are erased in terms of unknowns versus equations cf. [Zho06], but on the other hand lies their ambiguity. To be more precise, the magnetic flux density $\mathbf{B}$ remains unchanged if we change

$$\bar{\mathbf{A}} \to \mathbf{A} + \nabla\chi$$

that is if we added the gradient of an arbitrary differentiable scalar field $\chi : \Omega \times \mathcal{I} \to \mathbb{R}$ to $\mathbf{A}$. The same thing holds for the electric field strength $\mathbf{E}$ if we additionally replace

$$\bar{\boldsymbol{\varphi}} \to \boldsymbol{\varphi} - \frac{\partial}{\partial t}\chi$$

see for instance [Sch11]. In order to get rid of the ambiguity, we need additional so-called *gauge conditions*.

Gauge Classes Besides a variety of gauge conditions, the most famous ones are

1. *Coulomb gauge* condition
$$\nabla \cdot \mathbf{A} = 0$$
2. *Lorenz gauge* condition
$$\nabla \cdot \mathbf{A} + \varepsilon_0\mu_0\frac{\partial}{\partial t}\boldsymbol{\varphi} = 0$$

when assuming linear, homogeneous and isotropic materials, see [Jac99]. Hereby it is $\nu_0^{-1} = \mu_0 \approx 1.256\,637\,062\,12(19) \times 10^{-6}\,\mathrm{H\,m^{-1}}$ the permeability of vacuum and $\varepsilon_0 \approx 8.854\,187\,812\,8(13) \times 10^{-12}\,\mathrm{F\,m^{-1}}$ the permittivity of vacuum. Following the approach from [Bau12] we introduce the generalized gauge condition

$$\vartheta\varepsilon\nabla\frac{\partial}{\partial t}\boldsymbol{\varphi} + \zeta\nabla\big[\xi\nabla\cdot(\zeta\mathbf{A})\big] = 0 \tag{4.20}$$

with $\vartheta \in \mathbb{R}$ and artificial material rank-2 tensors $\zeta, \xi : \Omega \times \mathbb{R} \to \mathbb{R}^2$ whose choice is discussed for example in [CW02b] and [Cle05].
Setting a gauge freedom somehow corresponds to setting the ground level in circuits. For the forthcoming part of this treatise, we make use of the grad-type Lorenz gauge, that is (4.20) with $\vartheta = 1$.

4.4.3 Two System Formulations

In order to obtain a unique solution of the electromagnetic problem, involving conducting, isolating or semi-conducting media, we introduce two strategies.

First Variant The first variant persuades to solve a generalized grad-type gauge condition (4.20) together with the MA equation (4.18) exploiting (4.19), yielding for insulating and conducting media regions

$$\varepsilon\nabla\frac{\partial}{\partial t}\boldsymbol{\varphi}+\zeta\nabla\big[\xi\nabla\cdot(\zeta\mathbf{A})\big]=0, \tag{4.21a}$$

$$\nabla\times(\nu\nabla\times\mathbf{A})+\frac{\partial}{\partial t}\big[\varepsilon(\nabla\boldsymbol{\varphi}+\boldsymbol{\Pi})\big]+\sigma\big(\nabla\boldsymbol{\varphi}+\frac{\partial}{\partial t}\mathbf{A}\big)=0, \tag{4.21b}$$

$$\frac{\partial}{\partial t}\mathbf{A}-\boldsymbol{\Pi}=0. \tag{4.21c}$$

This system, as introduced in [Bau12], is in focus of this treatise. Note that we obtain the Gauss's law implicitly in a weak sense, i. e. taking the divergence of (4.21b) and defining $\frac{\partial}{\partial t}\boldsymbol{\rho}:=-\nabla\cdot\mathbf{J}$ in terms of the current continuity equation:

$$\begin{aligned}
\nabla\cdot\nabla\times(\nu\nabla\times\mathbf{A})+\nabla\cdot\big[\varepsilon\frac{\partial}{\partial t}\big(\nabla\boldsymbol{\varphi}+\boldsymbol{\Pi}\big)\big]+\nabla\cdot\Big[\sigma\big(\nabla\boldsymbol{\varphi}+\frac{\partial}{\partial t}\mathbf{A}\big)\Big]&=0\\
\nabla\cdot\frac{\partial}{\partial t}\big[\varepsilon(\nabla\boldsymbol{\varphi}+\boldsymbol{\Pi})\big]+\nabla\cdot\Big[\sigma\big(\nabla\boldsymbol{\varphi}+\frac{\partial}{\partial t}\mathbf{A}\big)\Big]&=0\\
\nabla\cdot\frac{\partial}{\partial t}\Big[\varepsilon\big(\nabla\boldsymbol{\varphi}+\frac{\partial}{\partial t}\mathbf{A}\big)\Big]-\nabla\cdot\mathbf{J}&=0\\
-\frac{\partial}{\partial t}\nabla\cdot\mathbf{D}-\nabla\cdot\mathbf{J}&=0\\
\frac{\partial}{\partial t}\nabla\cdot\mathbf{D}+\nabla\cdot\mathbf{J}&=0\\
\frac{\partial}{\partial t}\nabla\cdot\mathbf{D}&=\frac{\partial}{\partial t}\boldsymbol{\rho}
\end{aligned}$$

Since once satisfied, the Gauss's law is fulfilled all the time. For regions with semi-conducting media we have the system, cf. [Sch+19a],

$$\varepsilon\nabla\frac{\partial}{\partial t}\boldsymbol{\varphi}+\zeta\nabla\big[\xi\nabla\cdot(\zeta\mathbf{A})\big]=0, \tag{4.22a}$$

$$\nabla\cdot\mathbf{J}_{\mathbf{p}}+q\frac{\partial}{\partial t}\mathbf{p}+qR(\mathbf{p},\mathbf{n})=0, \tag{4.22b}$$

$$\nabla\cdot\mathbf{J}_{\mathbf{n}}-q\frac{\partial}{\partial t}\mathbf{n}-qR(\mathbf{p},\mathbf{n})=0, \tag{4.22c}$$

$$\mathbf{J}_{\mathbf{p}}+q\mu_{\mathrm{p}}\Big[\mathbf{p}\big(\nabla\boldsymbol{\varphi}+\frac{\partial}{\partial t}\mathbf{A}\big)+\frac{k\bar{T}}{q}\nabla\mathbf{p}\Big]=0, \tag{4.22d}$$

$$\mathbf{J}_{\mathbf{n}}+q\mu_{\mathrm{n}}\Big[\mathbf{n}\big(\nabla\boldsymbol{\varphi}+\frac{\partial}{\partial t}\mathbf{A}\big)-\frac{k\bar{T}}{q}\nabla\mathbf{n}\Big]=0, \tag{4.22e}$$

$$\nabla \times (\nu \nabla \times \mathbf{A}) + \frac{\partial}{\partial t}\big[\varepsilon(\nabla \boldsymbol{\varphi} + \boldsymbol{\Pi})\big] - \mathbf{J_p} - \mathbf{J_n} = 0, \tag{4.22f}$$

$$\frac{\partial}{\partial t}\mathbf{A} - \boldsymbol{\Pi} = 0. \tag{4.22g}$$

Second Variant Another strategy to solve the electromagnetic problem, which is similarly persuaded in the device modeler `DevEM` from the project partner *MAGWEL* [MAG16], is to solve the current continuity equation, incorporating the Gauss's law as $\boldsymbol{\rho}$, together with a modified Maxwell-Ampère equation and again the quasi-canonical momentum, see [Ban10; Che+11], reading

$$\nabla \cdot \big[\sigma(\nabla \boldsymbol{\varphi} + \boldsymbol{\Pi})\big] + \frac{\partial}{\partial t}\nabla \cdot \big[\varepsilon(\nabla \boldsymbol{\varphi} + \boldsymbol{\Pi})\big] = 0, \tag{4.23a}$$

$$\nabla \times (\nu \nabla \times \mathbf{A}) + \sigma(\nabla \boldsymbol{\varphi} + \boldsymbol{\Pi}) + \frac{\partial}{\partial t}\big[\varepsilon(\nabla \boldsymbol{\varphi} + \boldsymbol{\Pi})\big] - \nu_0 \nabla\big(\nabla \cdot \mathbf{A} + \mu_0 \varepsilon \frac{\partial}{\partial t}\boldsymbol{\varphi}\big) = 0, \tag{4.23b}$$

$$\frac{\partial}{\partial t}\mathbf{A} - \boldsymbol{\Pi} = 0. \tag{4.23c}$$

The modified MA equation (4.23b) equals the MA equation (4.21b) with a Lorenz-type gauge condition as a penalty term. Note that the electric charge density is now an output variable of Gauss's law and obtained by defining $\boldsymbol{\rho} := \nabla \cdot \mathbf{D} = \nabla \cdot [\varepsilon(\nabla \boldsymbol{\varphi} + \boldsymbol{\Pi})]$. With that, we immediately obtain charge conversation from (4.23a). It is also possible to recover the gauge condition from (4.23) by:

$$\begin{aligned}
\nabla \times (\nu \nabla \times \mathbf{A}) + \sigma(\nabla \boldsymbol{\varphi} + \boldsymbol{\Pi}) + \frac{\partial}{\partial t}\big[\varepsilon(\nabla \boldsymbol{\varphi} + \boldsymbol{\Pi})\big] - \nu_0 \nabla\big(\nabla \cdot \mathbf{A} + \mu_0 \varepsilon \frac{\partial}{\partial t}\boldsymbol{\varphi}\big) &= 0 \\
\nabla \cdot \big[\sigma(\nabla \boldsymbol{\varphi} + \boldsymbol{\Pi})\big] + \nabla \cdot \frac{\partial}{\partial t}\big[\varepsilon(\nabla \boldsymbol{\varphi} + \boldsymbol{\Pi})\big] - \nabla \cdot \Big[\nu_0 \nabla\big(\nabla \cdot \mathbf{A} + \mu_0 \varepsilon \frac{\partial}{\partial t}\boldsymbol{\varphi}\big)\Big] &= 0 \\
-\nabla \cdot \Big[\nu_0 \nabla\big(\nabla \cdot \mathbf{A} + \mu_0 \varepsilon \frac{\partial}{\partial t}\boldsymbol{\varphi}\big)\Big] &= 0 \\
\nabla^2\big(\nabla \cdot \mathbf{A} + \mu_0 \nabla \varepsilon \frac{\partial}{\partial t}\boldsymbol{\varphi}\big) &= 0
\end{aligned}$$

Hence, if the gauge condition is fulfilled on the boundary, it holds for the whole domain. A variant including semiconductors reads, in a scaled version,

$$\theta^{-1}\nabla \cdot \big[\frac{\varepsilon}{\varepsilon_0}(-\nabla \boldsymbol{\varphi} - \boldsymbol{\Pi})\big] - \mathbf{p} + \mathbf{n} - N_D^+ = 0, \tag{4.24a}$$

$$\nabla \cdot \mathbf{J_p} + q\frac{\partial}{\partial t}\mathbf{p} + qR(\mathbf{p}, \mathbf{n}) = 0, \tag{4.24b}$$

$$\nabla \cdot \mathbf{J_n} - q\frac{\partial}{\partial t}\mathbf{n} - qR(\mathbf{p}, \mathbf{n}) = 0, \tag{4.24c}$$

$$\mathbf{J_p} + q\mu_{\mathrm{p}}\Big[\mathbf{p}\big(\nabla \boldsymbol{\varphi} + \frac{\partial}{\partial t}\mathbf{A}\big) + \frac{k\bar{T}}{q}\nabla \mathbf{p}\Big] = 0, \tag{4.24d}$$

$$\mathbf{J_n} + q\mu_{\mathrm{n}}\Big[\mathbf{n}\big(\nabla \boldsymbol{\varphi} + \frac{\partial}{\partial t}\mathbf{A}\big) - \frac{k\bar{T}}{q}\nabla \mathbf{n}\Big] = 0, \tag{4.24e}$$

$$[\nabla \times (\nabla \times \mathbf{A}) - \nabla(\nabla \cdot \mathbf{A})] + K\frac{\partial}{\partial t}\big[\frac{\varepsilon}{\varepsilon_0}(\nabla \boldsymbol{\varphi} + \boldsymbol{\Pi})\big] \tag{4.24f}$$

$$-K\nabla\left(\frac{\varepsilon}{\varepsilon_0}\frac{\partial}{\partial t}\boldsymbol{\varphi}\right) + L\theta\sigma(\nabla\boldsymbol{\varphi} + \boldsymbol{\Pi}) - K\mathbf{J_p} - K\mathbf{J_n} = 0, \tag{4.24g}$$

$$\frac{\partial}{\partial t}\mathbf{A} - \boldsymbol{\Pi} = 0. \tag{4.24h}$$

where K and θ are scaling parameters, cf. [SCG14; Che+13].

Remark 4.7
Addressing the electromagnetic problem with the second variant, that is system (4.23), comes close to how it is actually done in several industrial solvers whereas the exact natures are certainly industrial secrets. Therefore and since it underwent already various analysis, the first variant is considered in this treatise, that is system (4.21).

4.5 Spatial Discretization using Finite Integration Technique

In this section we briefly explain the *finite integration technique* (FIT), originally introduced 1977 by Thomas Weiland [Wei77a]. The FIT is a discretization method developed to solve the inhomogeneous as well as the homogeneous MEs in a finite, three-dimensional, source-free region. Similar as the *finite difference time domain* (FDTD) method, developed by Yee [Yee66], it makes use of two staggered meshes. While originally introduced for a Cartesian mesh and homogeneous media, the method was further developed allowing now for more complex geometries and media, often inspired by enhancements for the FDTD method. Improvements include the consideration of anisotropic materials in [Krü00; Gut98] and various approaches to enhance the geometric approximation such as triangular, tetrahedral or adoptive meshes, subgridding techniques [Wei79; TW96; PCW03] as well as the *conformal finite integration technique* or the *nonorthogonal finite integration technique* [SW00; CW02a; Coo+06; SW98a].
Among other discretization techniques such as the *cell method* [Ton01] or the popular *finite element method* [LLC97; Bos98], the FIT with its recent enhancements is preferred by most simulation softwares. One of its outstanding features is that the deduced *Maxwell's grid equations* (MGEs) exactly solve the MEs up to discretization and that it has good conservation properties. The fields of application for electromagnetic problems reach from actuator simulation [FDW07], radio frequency simulation [Mat+19] and microwave simulation [MW07], to mention a few. Moreover, it is also used in the framework of scattering theory. For a deeper insight to applications we refer to FIT review in [Mar02].

First, it is given a possible discretization of a finite closed domain $\Omega \subset \mathbb{R}^3$, demonstrated for a brick shaped Ω with a rectilinear mesh. Followed by this, the concept of dual cells is brought to the audience. In order to define integral forms on these discretized geometrical structures, orientation is discussed. With these preliminaries integrals and discrete operators can be defined, which brings us the above mentioned Maxwell's grid equations. Contrary to [Bau12; Sch11], the here presented example of FIT for a brick shaped domain avoids the so-called ghost cells. This is how the implementation, with which the benchmarks in Chapter 7 are produced, is based on.

4.5.1 Discretization through Cell Complex

In order to spatially discretize the electromagnetic field problem, we first choose the domain of consideration $\Omega \subset \mathbb{R}^3$ fulfilling Assumption 4.5, independent of the method used. Then, for the domain of consideration, we define computational meshes by the so-called cell complexes, see [Cle98].

Let $\mathcal{X} = \{X_i\}_{i\in\mathbb{N}}$ be a countable partition of $\mathbb{R}^3$, such that Lebesgue measure $\lambda(X_i) > 0$ for all $i \in \mathbb{N}$. The available *geometrical objects* are

- volumes, which are by convention the closures $\bar{X}_i$;
- facets, as the intersections of two different volumes;
- links, as the intersections of at least two different facets;
- mesh points, as the intersections of at least two different links.

Respectively we refer to them as *p-cells*, for $p = 3, 2, 1, 0$, according to their geometrical dimensions.

In order to serve as computational mesh, we have to meet a few assumptions for the domain Ω and partition $\mathcal{X}$.

Assumption 4.8 (computational mesh)

(i) The domain of consideration $\Omega \subset \mathbb{R}^3$ is closed, bounded and simply connected.

(ii) For $\mathcal{X}$ there is a finite index set $\mathcal{N}_\mathcal{X}$ such that $\Omega = \{\bar{X}_i\}_{i\in\mathcal{N}_\mathcal{X}}$.

Notice that Assumption 4.8 implies Ω to be closed and that the p-cells can be mapped by finite index sets. Usually, we choose the volumes and domain Ω to be polytopes but the FIT in principle allows for all kind of cells that are homeomorphic to simplicial ones, cf. [CW01a]. For a spatial discretization using curved elements we refer to e. g. [Cia02]. This allows for the proper definition of a cell complex. Note that for solution theory, the domain of consideration Ω may fulfill stronger assumptions then the ones of Assumption 4.8.

Definition 4.9 (cell complex $\mathcal{G}$)
Given a domain Ω and partition $\mathcal{X}$ fulfilling Assumption 4.8, The p-cells obtained by $\{\bar{X}_i\}_{i\in\mathcal{N}_\mathcal{X}}$ define a so-called cell complex.

By convention, a cell complex is written as a set $\mathcal{G} = \mathcal{P} \cup \mathcal{L} \cup \mathcal{F} \cup \mathcal{V}$ with mesh points $\mathcal{P} = \{P(i) \mid i \in \mathcal{N}_\mathcal{P}\}$, links $\mathcal{L} = \{L(i) \mid i \in \mathcal{N}_\mathcal{L}\}$, facets $\mathcal{F} = \{F(i) \mid i \in \mathcal{N}_\mathcal{F}\}$ and volumes $\mathcal{V} = \{V(i) \mid i \in \mathcal{N}_\mathcal{V}\}$ whereby $\mathcal{N}_\mathcal{P}$, $\mathcal{N}_\mathcal{L}$, $\mathcal{N}_\mathcal{F}$ and $\mathcal{N}_\mathcal{V}$ are the according index sets. Note that the index sets may contain triples or other labels for the sake of construction which especially means that $\mathcal{N}_\mathcal{X}$ may differ from $\mathcal{N}_\mathcal{V}$.
The above definition of a cell complex generally allows all types of meshes, e. g. structured, unstructured, orthogonal, nonorthogonal, linear, curved etc. Whereas, structured meshes allow an easy identification of neighboring cells and hence construction of discrete operators,

indexed meshes require to store information about incident p-cells. Usually, the volumes are polyhedrons, such as cuboids or simplexes.

A typical choice for Ω are rectangular cuboids on which we can define *rectilinear* computational meshes. The advantage of such meshes is that they are structured meaning that the cell complexes can easily be constructed, see the following example.

Example 4.10 (rectilinear mesh on rectangular cuboid domain)
Let Ω be a rectangular cuboid such as in Figure 4.7. A cell complex $\mathcal{G}$ serving as a rectilinear computational mesh can be constructed as follows.

*Let N_x,$N_y, N_z \in \mathbb{N}$ be the desired number of mesh points, each greater than two, along each axis (*x-*,* y- *and* z*-axis) of Ω. That reasons to define the following index sets along the axes*

$$\begin{aligned} \mathcal{N}_x &:= \{1, \dots, N_x\}, \quad \mathcal{N}_x^- := \{1, \dots, N_x - 1\}, \\ \mathcal{N}_y &:= \{1, \dots, N_y\}, \quad \mathcal{N}_y^- := \{1, \dots, N_y - 1\}, \\ \mathcal{N}_z &:= \{1, \dots, N_z\}, \quad \mathcal{N}_z^- := \{1, \dots, N_z - 1\}. \end{aligned}$$

At this point we introduce the triple-index notation (i_x, i_y, i_z) according to the axes. Then, let the desired mesh point positions $(x_{i_x}, y_{i_y}, z_{i_z}) \in \Omega$ be arranged such that

$$x_1 < \dots < x_{i_x} < \dots < x_{N_x}, \quad y_1 < \dots < x_{i_y} < \dots < y_{N_y}, \quad z_1 < \dots < x_{i_z} < \dots < z_{N_z},$$

for $(i_x, i_y, i_z) \in \mathcal{N}_x \times \mathcal{N}_y \times \mathcal{N}_z$, see Figure 4.7, ensuring that

$$\Omega = \{[x_1, x_{N_x}] \times [y_1, y_{N_y}] \times [z_1, z_{N_z}]\}.$$

According to this indication, the following geometrical objects, see Figure 4.8, are naturally available for their axis dependent specific index sets:

- *mesh points, for $(i_x, i_y, i_z) \in \mathcal{N}_\mathcal{P} := \mathcal{N}_x \times \mathcal{N}_y \times \mathcal{N}_z$, defined by*

$$P(i_x, i_y, i_z) := \{x_{i_x} \times y_{i_y} \times z_{i_z}\},$$

- *links along* x*-axis, for $(i_x, i_y, i_z) \in \mathcal{N}_{\mathcal{L}_x} := \mathcal{N}_x^- \times \mathcal{N}_y \times \mathcal{N}_z$, defined by*

$$L(\mathrm{x}, i_x, i_y, i_z) := \{[x_{i_x}, x_{i_x+1}] \times y_{i_y} \times z_{i_z}\},$$

- *links along* y*-axis, for $(i_x, i_y, i_z) \in \mathcal{N}_{\mathcal{L}_y} := \mathcal{N}_x \times \mathcal{N}_y^- \times \mathcal{N}_z$, defined by*

$$L(\mathrm{y}, i_x, i_y, i_z) := \{x_{i_x} \times [y_{i_y}, y_{i_y+1}] \times z_{i_z}\},$$

- *links along* z*-axis, for $(i_x, i_y, i_z) \in \mathcal{N}_{\mathcal{L}_z} := \mathcal{N}_x \times \mathcal{N}_y \times \mathcal{N}_z^-$, defined by*

$$L(\mathrm{z}, i_x, i_y, i_z) := \{x_{i_x} \times y_{i_y} \times [z_{i_z}, z_{i_z+1}]\},$$

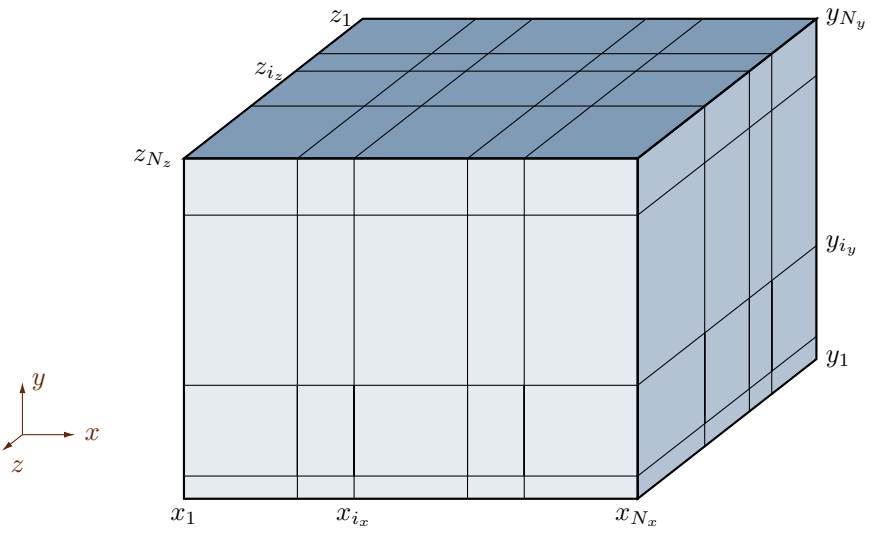


Figure 4.7: Example of a rectilinear mesh on a rectangular cuboid $\Omega \subset \mathbb{R}^3$. Grid points are obtained by combining each axis' values x_{i_x}, z_{i_z} and z_{i_z} (for $1 \leq i_x \leq N_y, 1 \leq i_y \leq N_y$ and $1 \leq i_z \leq N_z$).

- *facets orthogonal to* x*-axis, for* $(i_x, i_y, i_z) \in \mathcal{N}_{\mathcal{F}_x} := \mathcal{N}_x \times \mathcal{N}_y^- \times \mathcal{N}_z^-$*, defined by*

$$F(\mathrm{x}, i_x, i_y, i_z) := \{x_{i_x} \times [y_{i_y}, y_{i_y+1}] \times [z_{i_z}, z_{i_z+1}]\},$$

- *facets orthogonal to* y*-axis, for* $(i_x, i_y, i_z) \in \mathcal{N}_{\mathcal{F}_y} := \mathcal{N}_x^- \times \mathcal{N}_y \times \mathcal{N}_z^-$*, defined by*

$$F(\mathrm{y}, i_x, i_y, i_z) := \{[x_{i_x}, x_{i_x+1}] \times y_{i_y} \times [z_{i_z}, z_{i_z+1}]\},$$

- *facets orthogonal to* z*-axis, for* $(i_x, i_y, i_z) \in \mathcal{N}_{\mathcal{F}_z} := \mathcal{N}_x^- \times \mathcal{N}_y^- \times \mathcal{N}_z$*, defined by*

$$F(\mathrm{z}, i_x, i_y, i_z) := \{[x_{i_x}, x_{i_x+1}] \times [y_{i_y}, y_{i_y+1}] \times z_{i_z}\},$$

- *volumes, for* $(i_x, i_y, i_z) \in \mathcal{N}_{\mathcal{V}} := \mathcal{N}_x^- \times \mathcal{N}_y^- \times \mathcal{N}_z^-$*, defined by*

$$V(i_x, i_y, i_z) := \{[x_{i_x}, x_{i_x+1}] \times [y_{i_y}, y_{i_y+1}] \times [z_{i_z}, z_{i_z+1}]\}.$$

Note that not all geometrical structures are available for all triple indices, only for their, possible axis dependent, particular triple index sets. To be more precise, for a last point $(P(i_{N_X}, i_y, i_z))$ *in each axis direction, there is no further link* $(L(\mathrm{x}, i_{N_X}, i_y, i_z))$ *in this axis direction, since it would exceed* Ω *according to the same scheme. Moreover, the literals* x, y *and* z *of the first arguments for the links and facets are not to be confused with the coordinates, e. g.* $(x_{i_x}, y_{i_y}, z_{i_z}) \in \Omega$*. They serve as a label indicating the objects' alignment towards the axes. Merging the axis dependent index sets we obtain*

$$\mathcal{N}_{\mathcal{L}} := \{(w, i_x, i_y, i_z) \mid w \in \{\mathrm{x}, \mathrm{y}, \mathrm{z}\}, (i_x, i_y, i_z) \in \mathcal{N}_{\mathcal{L}_w}\},$$
$$\mathcal{N}_{\mathcal{F}} := \{(w, i_x, i_y, i_z) \mid w \in \{\mathrm{x}, \mathrm{y}, \mathrm{z}\}, (i_x, i_y, i_z) \in \mathcal{N}_{\mathcal{F}_w}\}.$$

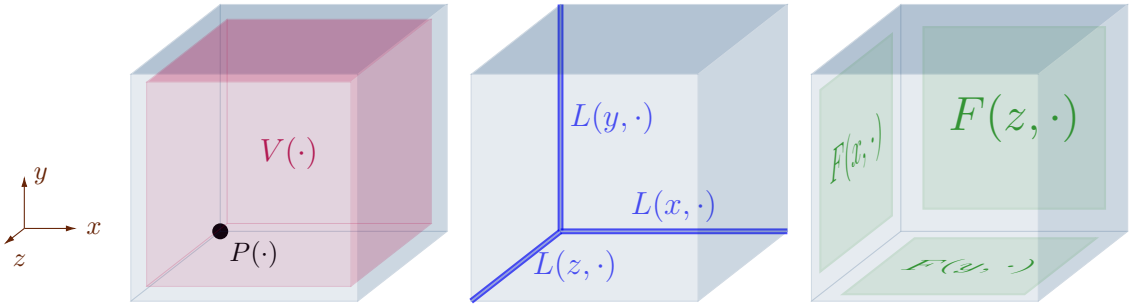


Figure 4.8: Labeling of Example 4.10. There is $P(\cdot)$ for mesh points, $V(\cdot)$ for volumes. Further, $L(\mathrm{x},\cdot), L(\mathrm{y},\cdot), L(\mathrm{z},\cdot)$ for links along and $F(\mathrm{x},\cdot), F(\mathrm{y},\cdot), F(\mathrm{z},\cdot)$ for facets orthogonal to the x-, y- and z-axis, respectively.

By making use of the above defined index sets,

$$\begin{aligned}
\mathcal{P} &:= \{P(i_x, i_y, i_z) \mid (i_x, i_y, i_z) \in \mathcal{N}_\mathcal{P}\}, \\
\mathcal{L} &:= \{L(w, i_x, i_y, i_z) \mid (w, i_x, i_y, i_z) \in \mathcal{N}_\mathcal{L}\}, \\
\mathcal{F} &:= \{F(w, i_x, i_y, i_z) \mid (w, i_x, i_y, i_z) \in \mathcal{N}_\mathcal{F}\}, \\
\mathcal{V} &:= \{V(i_x, i_y, i_z) \mid (i_x, i_y, i_z) \in \mathcal{N}_\mathcal{V}\}
\end{aligned}$$

are serving respectively as the p-cells sets for the cell complex $\mathcal{G}$.

4.5.2 Duality

One key property of the FIT is that the spatial discretization makes use of two cell complexes for Ω which are dual to each other by means that the computational meshes are staggered, see for example Figure 4.12. The duality of two cell complexes is given by the following definition.

Definition 4.11 (Dual cell complex $\widetilde{\mathcal{G}}$)
Let $\mathcal{G} = \mathcal{P} \cup \mathcal{L} \cup \mathcal{F} \cup \mathcal{V}$ and $\widetilde{\mathcal{G}} = \widetilde{\mathcal{P}} \cup \widetilde{\mathcal{L}} \cup \widetilde{\mathcal{F}} \cup \widetilde{\mathcal{V}}$ be two cell complexes of the same domain $\Omega \subset \mathbb{R}^3$. The cell complex $\widetilde{\mathcal{G}}$ is said to be a dual cell complex to the primal cell complex $\mathcal{G}$ if each p-cell of $\mathcal{G}$ is related to exactly one $(3-p)$-cell of $\widetilde{\mathcal{G}}$, in the following manner:

(i) Each primal mesh point is contained by exactly one dual volume, i. e.

$$\forall P(i) \in \mathcal{P} : \exists! \widetilde{V}(j) \in \widetilde{\mathcal{V}} : P \cap \widetilde{V} \neq \emptyset. \tag{4.25a}$$

(ii) Each primal link intersects exactly one inner dual facet, i. e.

$$\forall L(i) \in \mathcal{L} : \exists! \widetilde{F}(j) \in \widetilde{\mathcal{F}} : L \cap \widetilde{F} \neq \emptyset \text{ and } \widetilde{F} \not\subset \partial\Omega. \tag{4.25b}$$

(iii) Each primal face is intersected by exactly one inner dual link, i. e.

$$\forall F(i) \in \mathcal{F} : \exists! \widetilde{L}(j) \in \widetilde{\mathcal{L}} : F \cap \widetilde{L} \neq \emptyset \text{ and } \widetilde{L} \not\subset \partial\Omega. \tag{4.25c}$$

(iv) Each primal volume contains exactly one inner dual mesh point, i. e.

$$\forall V(i) \in \mathcal{V} : \exists! \widetilde{P}(j) \in \widetilde{\mathcal{P}} : V \cap \widetilde{P} \neq \emptyset \text{ and } \widetilde{P} \not\subset \partial\Omega. \tag{4.25d}$$

The dual index sets $\mathcal{N}_{\widetilde{\mathcal{V}}}$, $\mathcal{N}_{\widetilde{\mathcal{F}}}$, $\mathcal{N}_{\widetilde{\mathcal{L}}}$ and $\mathcal{N}_{\widetilde{\mathcal{P}}}$ are chosen such that (4.25) holds for $i = j$, respectively.

The geometrical structure is of importance leading to a few things that have to be discussed concerning these assumptions, see Remark 4.12.

Remark 4.12
Whereas in an infinite case ($\Omega = \mathbb{R}^3$) a perfect on-to-one relation of primal p-cells and dual $(3-p)$-cells is possible, the boundary of Ω somehow clips the dual cell complex $\widetilde{\mathcal{G}}$ leading to additional dual geometric objects at $\partial\Omega$ and disturbing the one-to-one dual to primal relation:

- *The outer facets, edges and mesh points of a dual volume belonging to a primal mesh point at the boundary, do not have a primal counterpart, see for instance Figures 4.9 and 4.12 of Example 4.10.*
- *From the other perspective: Primal volumes, facets or links at the boundary contain more then one dual mesh point, link or facets, respectively, when not excluding dual objects at the boundary. Therefore, the boundary exception for 1,2,3-cells, see (4.25b), (4.25c) and (4.25d) in Definition 4.11.*

Example 4.10 (continued)
One possible dual mesh $\widetilde{\mathcal{G}}$ to the primal mesh $\mathcal{G}$ can be obtained as follows. For

$$\begin{aligned} \mathcal{N}_{\widetilde{x}} &:= \{1, \dots, N_x + 1\}, & \mathcal{N}_{\widetilde{x}}^{-} &:= \{1, \dots, N_x\}, \\ \mathcal{N}_{\widetilde{y}} &:= \{1, \dots, N_y + 1\}, & \mathcal{N}_{\widetilde{y}}^{-} &:= \{1, \dots, N_y\}, \\ \mathcal{N}_{\widetilde{z}} &:= \{1, \dots, N_z + 1\}, & \mathcal{N}_{\widetilde{z}}^{-} &:= \{1, \dots, N_z\} \end{aligned}$$

we define the mesh points' coordinates per axis by

$$\begin{aligned} \widetilde{x}_1 &:= x_1, & \widetilde{x}_{N_x+1} &:= x_{N_x}, & \widetilde{x}_k &:= \frac{x_k + x_{k-1}}{2} \text{ for } k = 2, \dots, N_x, \\ \widetilde{y}_1 &:= y_1, & \widetilde{y}_{N_y+1} &:= y_{N_y}, & \widetilde{y}_k &:= \frac{y_k + y_{k-1}}{2} \text{ for } k = 2, \dots, N_y, \\ \widetilde{z}_1 &:= z_1, & \widetilde{z}_{N_z+1} &:= z_{N_z}, & \widetilde{z}_k &:= \frac{z_k + z_{k-1}}{2} \text{ for } k = 2, \dots, N_z. \end{aligned}$$

Using these coordinates, the dual geometric objects are deduced similar as before but this time with a shifted indexing scheme. In particular, we define

- *dual mesh points, for $(i_x, i_y, i_z) \in \mathcal{N}_{\widetilde{\mathcal{P}}} := \mathcal{N}_{\widetilde{x}} \times \mathcal{N}_{\widetilde{y}} \times \mathcal{N}_{\widetilde{z}}$ by*

$$\widetilde{P}(i_x - 1, i_y - 1, i_z - 1) := \{\widetilde{x}_{i_x} \times \widetilde{y}_{i_y} \times \widetilde{z}_{i_z}\},$$

- *dual links along* x*-axis, for* $(i_x, i_y, i_z) \in \mathcal{N}_{\widetilde{\mathcal{L}}_x} := \mathcal{N}_{\widetilde{x}}^- \times \mathcal{N}_{\widetilde{y}} \times \mathcal{N}_{\widetilde{z}}$ *by*
$$\widetilde{L}(\mathrm{x}, i_x, i_y - 1, i_z - 1) := \{[\widetilde{x}_{i_x}, \widetilde{x}_{i_x+1}] \times \widetilde{y}_{i_y} \times \widetilde{z}_{i_z}\}.$$
- *dual links along* y*-axis, for* $(i_x, i_y, i_z) \in \mathcal{N}_{\widetilde{\mathcal{L}}_y} := \mathcal{N}_{\widetilde{x}} \times \mathcal{N}_{\widetilde{y}}^- \times \mathcal{N}_{\widetilde{z}}$ *by*
$$\widetilde{L}(\mathrm{y}, i_x - 1, i_y, i_z - 1) := \{\widetilde{x}_{i_x} \times [\widetilde{y}_{i_y}, \widetilde{y}_{i_y+1}] \times \widetilde{z}_{i_z}\},$$
- *dual links along* z*-axis, for* $(i_x, i_y, i_z) \in \mathcal{N}_{\widetilde{\mathcal{L}}_z} := \mathcal{N}_{\widetilde{x}} \times \mathcal{N}_{\widetilde{y}} \times \mathcal{N}_{\widetilde{z}}^-$ *by*
$$\widetilde{L}(\mathrm{z}, i_x - 1, i_y - 1, i_z) := \{\widetilde{x}_{i_x} \times \widetilde{y}_{i_y} \times [\widetilde{z}_{i_z}, \widetilde{z}_{i_z+1}]\},$$
- *dual facets orthogonal to* x*-axis, for* $(i_x, i_y, i_z) \in \mathcal{N}_{\widetilde{\mathcal{F}}_x} := \mathcal{N}_{\widetilde{x}} \times \mathcal{N}_{\widetilde{y}}^- \times \mathcal{N}_{\widetilde{z}}^-$ *by*
$$\widetilde{F}(\mathrm{x}, i_x - 1, i_y, i_z) := \{\widetilde{x}_{i_x} \times [\widetilde{y}_{i_y}, \widetilde{y}_{i_y+1}] \times [\widetilde{z}_{i_z}, \widetilde{z}_{i_z+1}]\},$$
- *dual facets orthogonal to* y*-axis, for* $(i_x, i_y, i_z) \in \mathcal{N}_{\widetilde{\mathcal{F}}_y} := \mathcal{N}_{\widetilde{x}}^- \times \mathcal{N}_{\widetilde{y}} \times \mathcal{N}_{\widetilde{z}}^-$ *by*
$$\widetilde{F}(\mathrm{y}, i_x, i_y - 1, i_z) := \{[\widetilde{x}_{i_x}, \widetilde{x}_{i_x+1}] \times \widetilde{y}_{i_y} \times [\widetilde{z}_{i_z}, \widetilde{z}_{i_z+1}]\},$$
- *dual facets orthogonal to* z*-axis, for* $(i_x, i_y, i_z) \in \mathcal{N}_{\widetilde{\mathcal{F}}_z} := \mathcal{N}_{\widetilde{x}}^- \times \mathcal{N}_{\widetilde{y}}^- \times \mathcal{N}_{\widetilde{z}}$ *by*
$$\widetilde{F}(\mathrm{z}, i_x, i_y, i_z - 1) := \{[\widetilde{x}_{i_x}, \widetilde{x}_{i_x+1}] \times [\widetilde{y}_{i_y}, \widetilde{y}_{i_y+1}] \times \widetilde{z}_{i_z}\},$$
- *dual volumes for* $(i_x, i_y, i_z) \in \mathcal{N}_{\widetilde{\mathcal{V}}} := \mathcal{N}_{\widetilde{x}}^- \times \mathcal{N}_{\widetilde{y}}^- \times \mathcal{N}_{\widetilde{z}}^-$ *by*
$$\widetilde{V}(i_x, i_y, i_z) := \{[\widetilde{x}_{i_x}, \widetilde{x}_{i_x+1}] \times [\widetilde{y}_{i_y}, \widetilde{y}_{i_y+1}] \times [\widetilde{z}_{i_z}, \widetilde{z}_{i_z+1}]\}.$$

Further, we have the cumulated dual links' and facets' index sets
$$\mathcal{N}_{\widetilde{\mathcal{L}}} := \left\{(w, i_x, i_y, i_z) \mid w \in \{\mathrm{x}, \mathrm{y}, \mathrm{z}\}, (i_x, i_y, i_z) \in \mathcal{N}_{\widetilde{\mathcal{L}}_w}\right\},$$
$$\mathcal{N}_{\widetilde{\mathcal{F}}} := \left\{(w, i_x, i_y, i_z) \mid w \in \{\mathrm{x}, \mathrm{y}, \mathrm{z}\}, (i_x, i_y, i_z) \in \mathcal{N}_{\widetilde{\mathcal{F}}_w}\right\}.$$
It is easy to verify that, for the same index, which may also include the axis label, the according p-cells comply the one-to-one relation introduced in (4.25). To be more precise, each primal mesh point $P(i)$, for $i \in \mathcal{N}_{\mathcal{P}} \subset \mathcal{N}_{\widetilde{\mathcal{P}}}$, is exclusively contained in the dual volume $\widetilde{V}(i)$, see e. g. Figure 4.9. Further, $L(i)$ intersects $\widetilde{F}(i)$, for $i \in \mathcal{N}_{\mathcal{L}} \subset \mathcal{N}_{\widetilde{\mathcal{F}}}$, and $F(i)$ is intersected by $\widetilde{L}(i)$, for $i \in \mathcal{N}_{\mathcal{F}} \subset \mathcal{N}_{\widetilde{\mathcal{L}}}$, see respectively Figures 4.10 and 4.11. Finally, every primal volume $V(i)$ contains $\widetilde{P}(i)$, for $i \in \mathcal{N}_{\mathcal{V}} \subset \mathcal{N}_{\widetilde{\mathcal{P}}}$, see again Figure 4.8. Notice that the remaining dual objects are part of the boundary, that is
$$\widetilde{F}(i) \subset \partial\Omega, \text{ for } i \in \mathcal{N}_{\widetilde{\mathcal{F}}} \backslash \mathcal{N}_{\mathcal{L}},$$
$$\widetilde{L}(i) \subset \partial\Omega, \text{ for } i \in \mathcal{N}_{\widetilde{\mathcal{L}}} \backslash \mathcal{N}_{\mathcal{F}},$$
$$\widetilde{P}(i) \subset \partial\Omega, \text{ for } i \in \mathcal{N}_{\widetilde{\mathcal{P}}} \backslash \mathcal{N}_{\mathcal{V}}.$$
Moreover, due to the rectilinear structure, primal links are orthogonal to dual facets and vice versa which is why we speak of orthogonal staggered cell complexes.

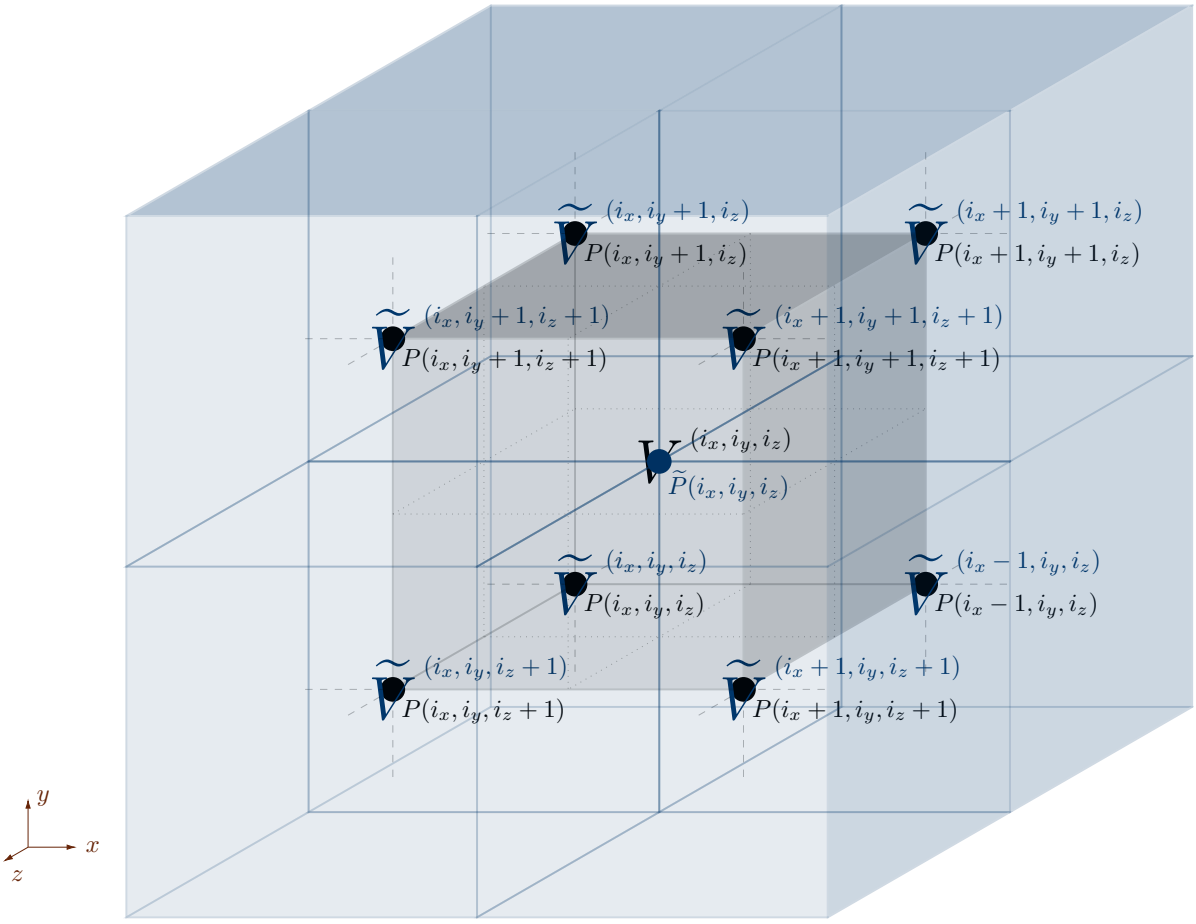


Figure 4.9: Primal volumes and mesh points (black) with their dual mesh points' and volumes' (blue) respective counterpart, of Example 4.10. At the boundary, $\widetilde{V}(i)$ is clipped and does not exceed Ω which is the case if $P(i) \subset \partial\Omega$.

4.5.3 Orientation

The p-cells, for $p \in \{1, 2, 3\}$, of the cell complexes $\mathcal{G}$ and $\widetilde{\mathcal{G}}$, serve as domains for integration. Hence, we need a concept of orientation to define integral quantities on them such as for the tangential and vector fields. As of Stokes' or Gauss's theorem, these domains are linked to their $(p-1)$-surfaces which is why we need also an orientation for them. Notice that it is one property of the cell complexes, that each p-cell surface is an accumulation of $(p-1)$-cells.

Remark 4.13
Let $\mathcal{G}$ be a cell complex. By construction, each $(p-1)$-surface of a p-cell of $\mathcal{G}$ can be expressed as a composition of $(p-1)$-cells, i. e. for $V \in \mathcal{V}$, $F \in \mathcal{F}$ and $L \in \mathcal{L}$ it holds

$$\partial V = \bigcup_{F \in \mathcal{F}|_V} F, \quad \partial F = \bigcup_{L \in \mathcal{L}|_F} L, \quad \text{and} \quad \partial L = \bigcup_{P \in \mathcal{P}|_L} P,$$

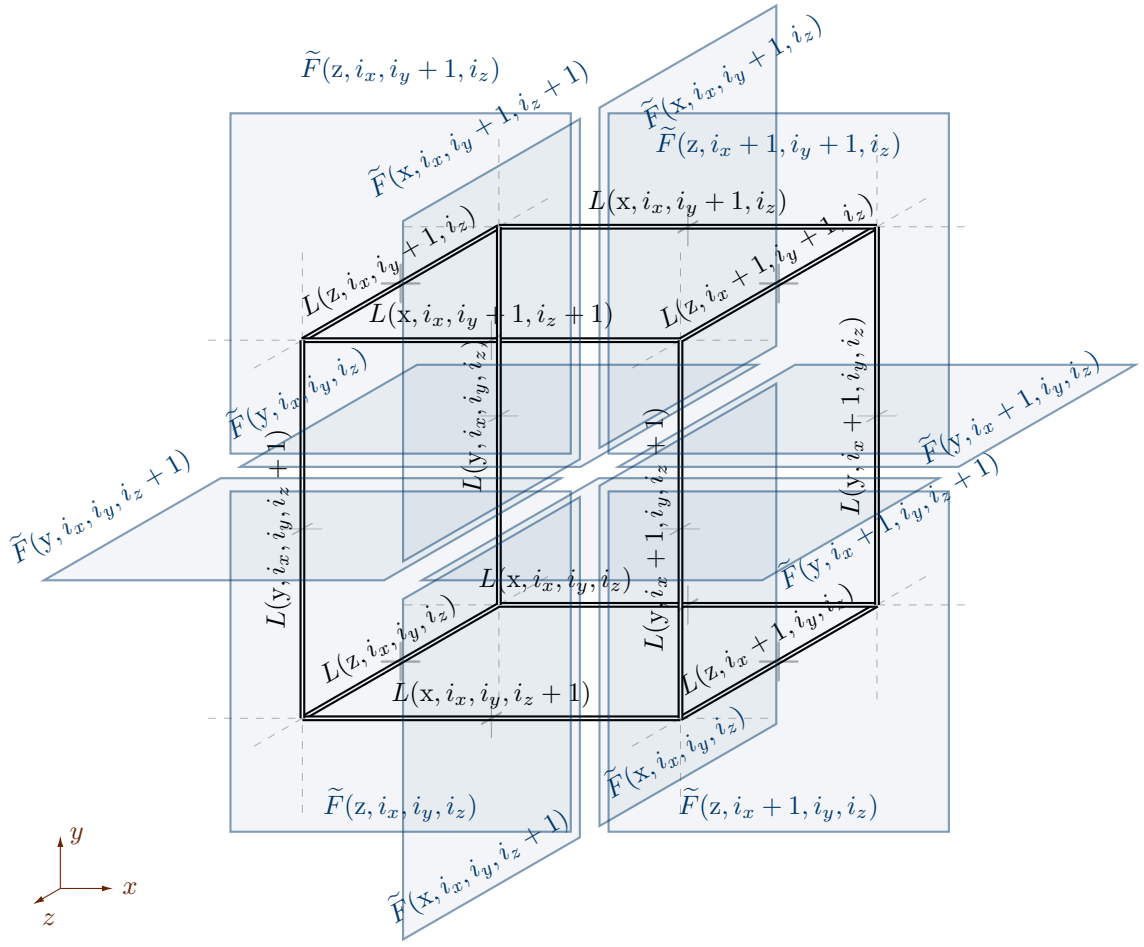


Figure 4.10: Primal links (black) with their dual facets' (blue) respective counterpart, of Example 4.10. At the boundary $\widetilde{F}(i)$ is clipped and does not exceed Ω which is the case if $L(i) \subset \partial\Omega$.

respectively, where the $(p-1)$*-cells of a corresponding* p*-cell are denoted with*

$$\mathcal{F}|_V = \{F \in \mathcal{F} \mid F \subset V\}, \quad \mathcal{L}|_F = \{L \in \mathcal{L} \mid L \subset F\} \quad \textit{and} \quad \mathcal{P}|_L = \{P \in \mathcal{P} \mid P \subset L\}.$$

Note that the orientation of the p-cell boundaries is following the convention that the unit normal field of ∂V shall be pointed outwards the volume, the tangential field if of ∂F is counterclockwise viewed from top-down the facet's normal field and source and sink of ∂L are oriented from start to end of the link's tangential field. The orientations' match or mismatch information of $(p-1)$-cells and their superseding p-cells' boundaries are stored in the so-called *window functions*.

Definition 4.14 (window functions)
Let $\mathcal{G}$ *be a cell complex such that each mesh point, link and facet has a fixed orientation. The*

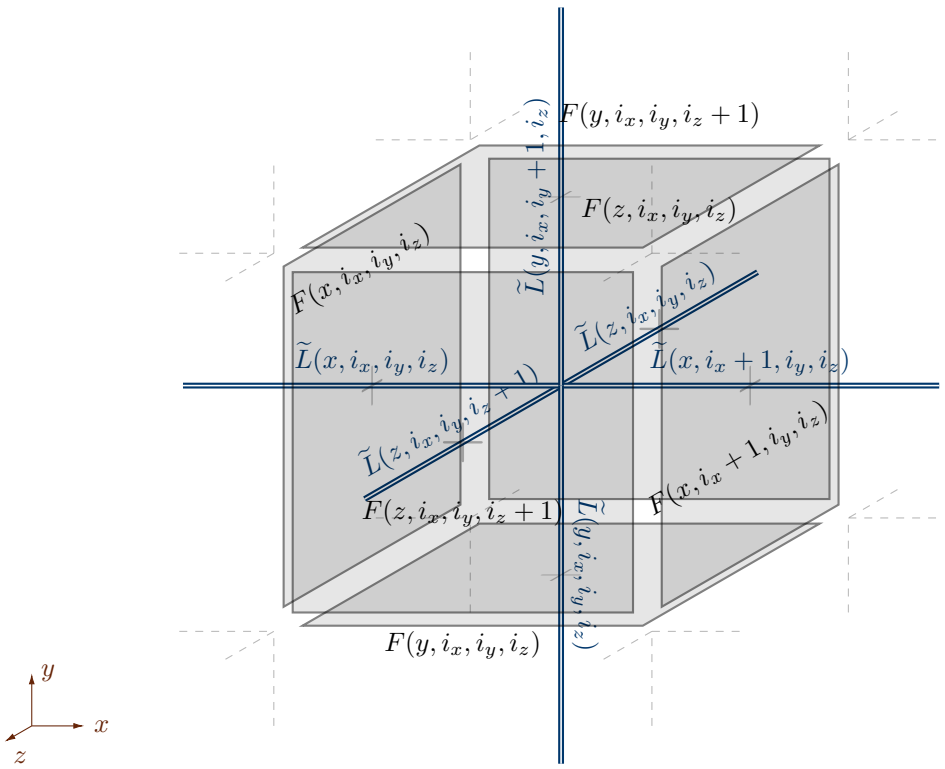


Figure 4.11: Primal facets (black) with their dual links' (blue) respective counterpart, of Example 4.10. At the boundary $\widetilde{L}(i)$ is clipped and does not exceed Ω which is the case if $F(i) \subset \partial\Omega$.

functions

$$\zeta^{[1]} : \mathcal{P} \times \mathcal{L} \to \{-1, 0, 1\}, \quad \zeta^{[2]} : \mathcal{L} \times \mathcal{F} \to \{-1, 0, 1\}, \quad \text{and} \quad \zeta^{[3]} : \mathcal{F} \times \mathcal{V} \to \{-1, 0, 1\}$$

with

$$\zeta^{[p]}(G, H) := \begin{cases} 1 & \text{if } G \text{ and } \partial H \cap G \text{ have equal orientations,} \\ -1 & \text{if } G \text{ and } \partial H \cap G \text{ have opposite orientations,} \\ 0 & \text{if } G \not\subset \partial H. \end{cases}$$

for $p \in \{1, 2, 3\}$, are called the window functions of $\mathcal{G}$.

For the sake of simplicity make a few assumptions on the cell complexes.

Assumption 4.15 (orthogonal staggered)
The primal and dual cell complexes $\mathcal{G}$ and $\widetilde{\mathcal{G}}$ are chosen such that

(i) Primal links are orthogonal to their dual facets counterpart, and vice versa.

(ii) Primal and dual counterparts have the same orientation.

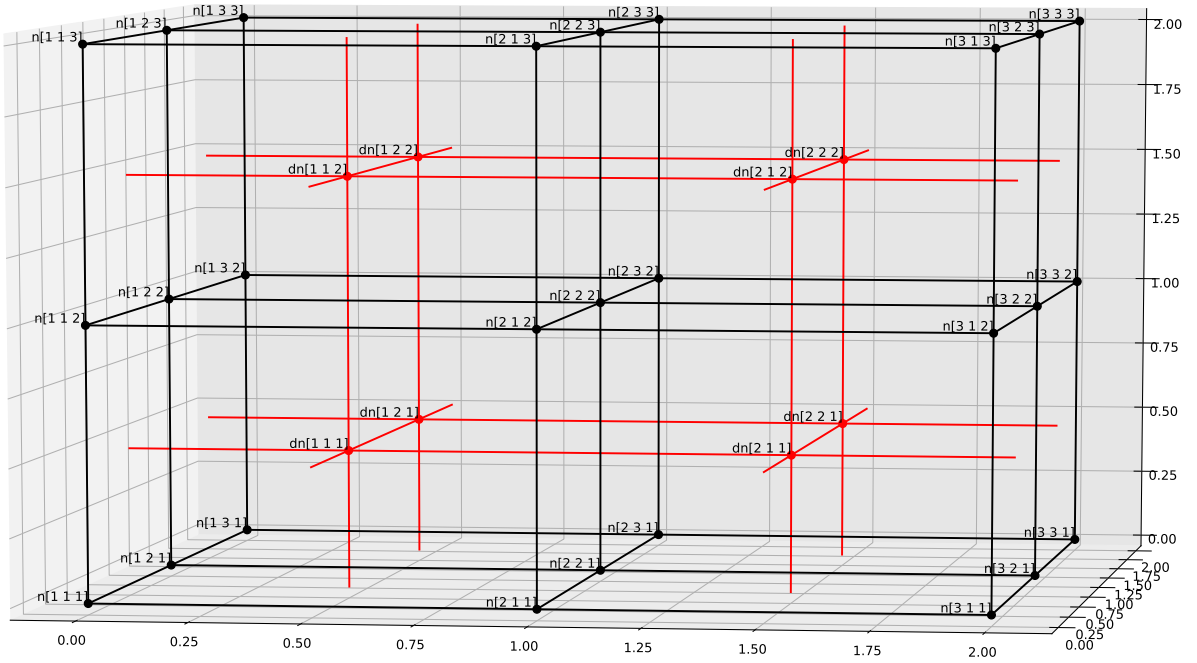


Figure 4.12: Example of staggered primal (black) and dual (red) cell complex with boundary. Red dots and lines refer to dual 0- and 1-cells of the according internal sets.

Assumption 4.15 is in general not necessary for making use of the FIT if carefully considered in the discrete equations' derivation, see for instance the already mentioned nonorthogonal finite integration technique in [SW00], but allows for an easier introduction and understanding.

Example 4.10 (continued)
Albeit it can be arbitrary, we introduce the following pragmatic orientation scheme for the here considered rectilinear cell complexes $\mathcal{G}$ and $\widetilde{\mathcal{G}}$. For each primal link $L(w, i_x, i_y.i_z) \in \mathcal{L}$ and primal facet $F(w, i_x, i_y.i_z) \in \mathcal{F}$, we choose respectively the unit tangent and unit normal fields' direction to be oriented along the w-axis ($w \in \{\mathrm{x}, \mathrm{y}, \mathrm{z}\}$), see Figure (4.13). Further, primal mesh points are identified as sources. Analogously, the same is done for all dual mesh points, links and facets. This are the orientations assigned to the $(p-1)$-cells for $p \in \{1, 2, 3\}$. With this convention, the unit tangent field of each, primal and dual link, equals the unit normal field of the dual and primal counterpart. To give an example, the unit normal field of facet $F(\mathrm{x}, 1, 1, 1)$ and the unit tangent field of the according dual link $\widetilde{L}(\mathrm{x}, 1, 1, 1)$ are both valued $(1, 0, 0)$. Together with the previous observation of orthogonality, Assumption 4.15 is fulfilled.

4.5.4 Maxwell's Grid Equations

As a next step to a spatially discretized electromagnetic problem, Maxwell's equations in their integral form and the constitutive relations are mapped to the cell complexes $\mathcal{G}$ and $\widetilde{\mathcal{G}}$. Instead of assigning field strengths, fluxes or densities to a certain point, new integral state quantities, i. e. electromagnetic quantities integrated over p-cells, are introduced, see [Wei77a].

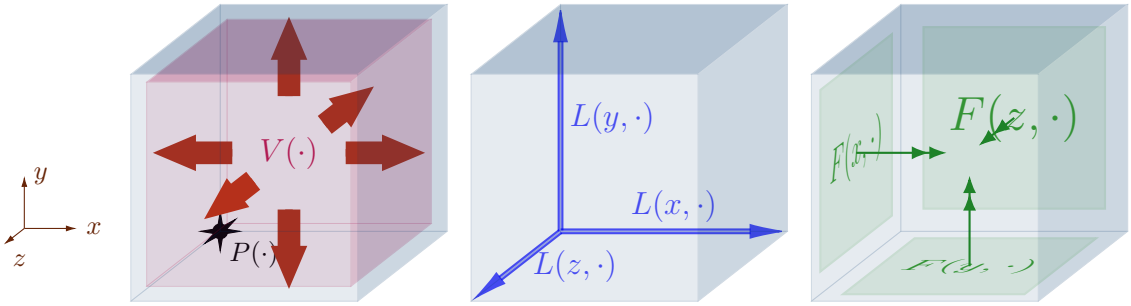


Figure 4.13: This figure shows the orientation of th p-cells of $\mathcal{G}$ of Example 4.10. P points inwords, V points outwords, L_x, L_y, L_Z parallely follow their axes and F_x, F_y, F_z are pointing in direction of their orthogonal axes.

We start by recalling MEs in integral formulation (4.2)

$$\begin{aligned}
\oiint_{\partial V} \mathbf{D}(s,t) \cdot \mathrm{d}s &= \iiint_V \boldsymbol{\rho}(r,t)\,\mathrm{d}r,\\
\oiint_{\partial V} \mathbf{B}(s,t) \cdot \mathrm{d}s &= 0,\\
\oint_{\partial S} \mathbf{E}(l,t) \cdot \mathrm{d}l &= -\iint_S \frac{\partial}{\partial t}\mathbf{B}(s,t) \cdot \mathrm{d}s,\\
\oint_{\partial S} \mathbf{H}(l,t) \cdot \mathrm{d}l &= \iint_S \left(\mathbf{J}(s,t) + \frac{\partial}{\partial t}\mathbf{D}(s,t)\right) \cdot \mathrm{d}s.
\end{aligned}$$

The encountered electromagnetic quantities are integrated over either volumes, i. e. V, surfaces, i. e. ∂V and S or contours, i. e. ∂S. As they hold for arbitrary piecewise smooth p-cells, let them be the p-cells of $\mathcal{G}$ and $\widetilde{\mathcal{G}}$. Together with the observations in Remark 4.13 and by making use of the window functions $\zeta^{[p]}$ of $\mathcal{G}$ and $\widetilde{\zeta}^{[p]}$ of $\widetilde{\mathcal{G}}$, the so-called *Maxwell's grid equations* (MGEs) on the cell complexes read

- Gauss's law

$$\sum_{\widetilde{F} \in \widetilde{\mathcal{F}}|_{\widetilde{V}}} \widetilde{\zeta}^{[3]}(\widetilde{F}, \widetilde{V}) \iint_{\widetilde{F}} \mathbf{D}(s,t) \cdot \mathrm{d}s = \iiint_{\widetilde{V}} \boldsymbol{\rho}(r,t)\,\mathrm{d}r, \quad \forall \widetilde{V} \in \widetilde{\mathcal{V}} \tag{4.26a}$$

- Gauss's law for magnetism

$$\sum_{F \in \mathcal{F}|_V} \zeta^{[3]}(F, V) \iint_F \mathbf{B}(s,t) \cdot \mathrm{d}s = 0, \quad \forall V \in \mathcal{V} \tag{4.26b}$$

- Maxwell-Faraday's law

$$\sum_{L \in \mathcal{L}|_F} \zeta^{[2]}(L, F) \int_L \mathbf{E}(l,t) \cdot \mathrm{d}l = -\iint_F \frac{\partial}{\partial t}\mathbf{B}(s,t) \cdot \mathrm{d}s, \quad \forall F \in \mathcal{F} \tag{4.26c}$$

- Maxwell-Ampère's law

$$\sum_{\widetilde{L}\in\widetilde{\mathcal{L}}|_{\widetilde{F}}} \widetilde{\zeta}^{[2]}(\widetilde{L},\widetilde{F}) \int_{\widetilde{L}} \mathbf{H}(l,t)\cdot \mathrm{d}l = \iint_{\widetilde{F}} \left(\mathbf{J}(s,t) + \frac{\partial}{\partial t}\mathbf{D}(s,t)\right)\cdot \mathrm{d}s, \quad \forall \widetilde{F}\in\widetilde{\mathcal{F}} \tag{4.26d}$$

Note that each equation of MEs can be applied to any cell complex, but the FIT splits them up on $\mathcal{G}$ and its dual $\widetilde{\mathcal{G}}$ for the number of resulting integral quantities. Further, the primal and dual meshes' role was fixed by convention and could be flipped.

Remark 4.16
In the following we use vectors (or vector function) and matrices which are defined in terms of indexed families. This is to be understood as follows. Let $\mathcal{N}_1$ and $\mathcal{N}_2$ be a finite index sets of unique elements, called labels or indices, and $\mathbb{K}$ be a field. With $v\in\mathbb{K}^{\mathcal{N}_1}$ and $M\in\mathbb{K}^{\mathcal{N}_2\times\mathcal{N}_1}$ we denote, similar to vectors and matrices, families defined by their components

$$v := (v_i)_{i\in\mathcal{N}_1}, \qquad M := (M_{i,j})_{i\in\mathcal{N}_2, j\in\mathcal{N}_1}.$$

Besides the fact that their components' order is not fixed, the vector and matrix concepts apply straight forward to v and M, e. g. the product Mv yields another indexed family $w\in\mathbb{R}^{\mathcal{N}_2}$ with $w_i = \sum_{j\in\mathcal{N}_1} M_{i,j}v_j$.

According to the indexing of all p-cells, the mesh quantities are introduced as the integrals

$$\widehat{\mathbf{e}}_i(t) := \int_{L(i)} \mathbf{E}(l,t)\cdot \mathrm{d}l \qquad \forall L(i)\in\mathcal{L}, \tag{4.27a}$$

$$\widehat{\widehat{\mathbf{b}}}_i(t) := \iint_{F(i)} \mathbf{B}(s,t)\cdot \mathrm{d}s \qquad \forall F(i)\in\mathcal{F}, \tag{4.27b}$$

$$\widehat{\mathbf{h}}^{\mathrm{exa}}_i(t) := \int_{\widetilde{L}(i)} \mathbf{H}(l,t)\cdot \mathrm{d}l \qquad \forall \widetilde{L}(i)\in\widetilde{\mathcal{L}}, \tag{4.27c}$$

$$\widehat{\widehat{\mathbf{d}}}^{\mathrm{exa}}_i(t) := \iint_{\widetilde{F}(i)} \mathbf{D}(s,t)\cdot \mathrm{d}s \qquad \forall \widetilde{F}(i)\in\widetilde{\mathcal{F}}, \tag{4.27d}$$

$$\widehat{\widehat{\mathbf{j}}}^{\mathrm{exa}}_i(t) := \iint_{\widetilde{F}(i)} \mathbf{J}(s,t)\cdot \mathrm{d}s \qquad \forall \widetilde{F}(i)\in\widetilde{\mathcal{F}}, \tag{4.27e}$$

$$\widehat{\widehat{\mathbf{q}}}_i(t) := \iiint_{\widetilde{V}(i)} \boldsymbol{\rho}(r,t)\,\mathrm{d}r \qquad \forall \widetilde{V}(i)\in\widetilde{\mathcal{V}}, \tag{4.27f}$$

with $\widehat{\mathbf{e}}:\mathcal{I}\to\mathbb{R}^{\mathcal{N}_\mathcal{L}}$ the *electric mesh voltages* and $\widehat{\widehat{\mathbf{b}}}:\mathcal{I}\to\mathbb{R}^{\mathcal{N}_\mathcal{F}}$ the *magnetic mesh fluxes*. Since $\mathbf{E}$ and $\mathbf{B}$ are related by Maxwell-Faraday's law (4.2c), $\widehat{\mathbf{e}}$ and $\widehat{\widehat{\mathbf{b}}}$ are assigned to p-cells of the same cell complex, in this case the primal cell complex $\mathcal{G}$. On the dual p-cells we allocate $\widehat{\mathbf{h}}^{\mathrm{exa}}:\mathcal{I}\to\mathbb{R}^{\mathcal{N}_{\widetilde{\mathcal{L}}}}$ the *magnetic mesh voltages*, $\widehat{\widehat{\mathbf{d}}}^{\mathrm{exa}}:\mathcal{I}\to\mathbb{R}^{\mathcal{N}_{\widetilde{\mathcal{F}}}}$ the *electric mesh fluxes*, $\widehat{\widehat{\mathbf{j}}}^{\mathrm{exa}}:\mathcal{I}\to\mathbb{R}^{\mathcal{N}_{\widetilde{\mathcal{F}}}}$ the *electric mesh currents* and $\widehat{\widehat{\mathbf{q}}}:\mathcal{I}\to\mathbb{R}^{\mathcal{N}_{\widetilde{\mathcal{V}}}}$ the *distribution of charges* due to their relation by Maxwell-Ampère's law (4.2d).
For a better orientation, $\widehat{\mathbf{e}}$ and $\widehat{\mathbf{h}}^{\mathrm{exa}}$ are decorated with one arc "$\frown$" since they express line integrals. According to expectations the double arcs "$\widehat{\ }$" of $\widehat{\widehat{\mathbf{b}}}$, $\widehat{\widehat{\mathbf{d}}}^{\mathrm{exa}}$ and $\widehat{\widehat{\mathbf{j}}}^{\mathrm{exa}}$ indicate quantities

representing surface integrals, as well as with "$\widehat{\widehat{\ }}$" of $\widehat{\widehat{\mathbf{q}}}$ we denote volume integrals, see [Bos88]. Note that since we have chosen an orientation, these mesh quantities (4.27) are well defined. From the window functions we can deduce the following discrete operators on $\mathcal{G}$

$$\begin{aligned}
G \in \{-1,0,1\}^{\mathcal{N}_\mathcal{L}\times\mathcal{N}_\mathcal{P}} &: G_{ij} = \zeta^{[1]}(P(j),L(i)), && \text{for } i\in\mathcal{N}_\mathcal{L}, j\in\mathcal{N}_\mathcal{P},\\
C \in \{-1,0,1\}^{\mathcal{N}_\mathcal{F}\times\mathcal{N}_\mathcal{L}} &: C_{ij} = \zeta^{[2]}(L(j),F(i)), && \text{for } i\in\mathcal{N}_\mathcal{F}, j\in\mathcal{N}_\mathcal{L},\\
S \in \{-1,0,1\}^{\mathcal{N}_\mathcal{V}\times\mathcal{N}_\mathcal{F}} &: S_{ij} = \zeta^{[3]}(F(j),V(i)), && \text{for } i\in\mathcal{N}_\mathcal{V}, j\in\mathcal{N}_\mathcal{F}
\end{aligned}$$

and analogously the discrete operators on $\widetilde{\mathcal{G}}$

$$\begin{aligned}
\widetilde{G}^{\text{exa}} \in \{-1,0,1\}^{\mathcal{N}_{\widetilde{\mathcal{L}}}\times\mathcal{N}_{\widetilde{\mathcal{P}}}} &: \widetilde{G}^{\text{exa}}_{ij} = \widetilde{\zeta}^{[1]}(\widetilde{P}(j),\widetilde{L}(i)), && \text{for } i\in\mathcal{N}_{\widetilde{\mathcal{L}}}, j\in\mathcal{N}_{\widetilde{\mathcal{P}}},\\
\widetilde{C}^{\text{exa}} \in \{-1,0,1\}^{\mathcal{N}_{\widetilde{\mathcal{F}}}\times\mathcal{N}_{\widetilde{\mathcal{L}}}} &: \widetilde{C}^{\text{exa}}_{ij} = \widetilde{\zeta}^{[2]}(\widetilde{L}(j),\widetilde{F}(i)), && \text{for } i\in\mathcal{N}_{\widetilde{\mathcal{F}}}, j\in\mathcal{N}_{\widetilde{\mathcal{L}}},\\
\widetilde{S}^{\text{exa}} \in \{-1,0,1\}^{\mathcal{N}_{\widetilde{\mathcal{V}}}\times\mathcal{N}_{\widetilde{\mathcal{F}}}} &: \widetilde{S}^{\text{exa}}_{ij} = \widetilde{\zeta}^{[3]}(\widetilde{F}(j),\widetilde{V}(i)), && \text{for } i\in\mathcal{N}_{\widetilde{\mathcal{V}}}, j\in\mathcal{N}_{\widetilde{\mathcal{F}}}.
\end{aligned}$$

With these discrete operators and the introduction of the mesh quantities (4.27), we can shorthandedly write MGEs (4.26) as

$$\widetilde{S}^{\text{exa}}\widehat{\widehat{\mathbf{d}}}^{\text{exa}} = \widehat{\widehat{\mathbf{q}}}, \tag{4.28a}$$

$$S\widehat{\widehat{\mathbf{b}}} = 0, \tag{4.28b}$$

$$C\widehat{\mathbf{e}} = -\frac{\partial}{\partial t}\widehat{\widehat{\mathbf{b}}}, \tag{4.28c}$$

$$\widetilde{C}^{\text{exa}}\widehat{\mathbf{h}}^{\text{exa}} = \widehat{\widehat{\mathbf{j}}}^{\text{exa}} + \frac{\partial}{\partial t}\widehat{\widehat{\mathbf{d}}}^{\text{exa}}. \tag{4.28d}$$

Remark 4.17
So far no approximation happened in the four MGEs (4.28) in scope of the p-cells since the ME are applied directly to the cell complexes and obtain only topological information, see [CW01b].

Note, the discrete operators mirror the continuous operator identities, i. e. $\nabla\cdot\nabla\times F = 0$ and $\nabla\times\nabla\psi = 0$ for arbitrary but sufficiently smooth three dimensional vector field F and scalar field ψ, in that

$$CG = 0, \qquad \widetilde{C}^{\text{exa}}\widetilde{G}^{\text{exa}} = 0, \tag{4.29a}$$

$$SC = 0, \qquad \widetilde{S}^{\text{exa}}\widetilde{C}^{\text{exa}} = 0. \tag{4.29b}$$

The discrete pendant of total current flow through a surface (4.15) translates to dual facets that is, for $\widetilde{F}(i)\in\widetilde{\mathcal{F}}$, the total current flow through that surface is given by

$$\left(\widehat{\widehat{\mathbf{j}}}^{\text{exa}} + \frac{\partial}{\partial t}\widehat{\widehat{\mathbf{d}}}^{\text{exa}}\right)_i = \left(\widetilde{C}^{\text{exa}}\widehat{\mathbf{h}}^{\text{exa}}\right)_i \tag{4.30}$$

as of (4.28d). As a next step of the FIT, a pure one-to-one relation of the primal and dual cells is established by separating the dual boundary's p-cells from the quantities and operators. The

absence of dual p-cells at the border is compensated in a subsequent step by the introduction of boundary conditions. Consider the new operators based upon primal index sets

$$\widetilde{G} := (\widetilde{G}_{ij}^{\text{exa}})_{i\in\mathcal{N}_{\mathcal{F}},j\in\mathcal{N}_{\mathcal{V}}}, \qquad \widehat{\mathbf{h}} := (\widehat{\mathbf{h}}_i^{\text{exa}})_{i\in\mathcal{N}_{\mathcal{F}}},$$
$$\widetilde{S} := (\widetilde{S}_{ij}^{\text{exa}})_{i\in\mathcal{N}_{\mathcal{P}},j\in\mathcal{N}_{\mathcal{L}}}, \qquad \widehat{\widetilde{\mathbf{d}}} := (\widehat{\widetilde{\mathbf{d}}}_i^{\text{exa}})_{i\in\mathcal{N}_{\mathcal{L}}},$$
$$\widetilde{C} := (\widetilde{C}_{ij}^{\text{exa}})_{i\in\mathcal{N}_{\mathcal{L}},j\in\mathcal{N}_{\mathcal{F}}}, \qquad \widehat{\widetilde{\mathbf{j}}} := (\widehat{\widetilde{\mathbf{j}}}_i^{\text{exa}})_{i\in\mathcal{N}_{\mathcal{L}}}.$$

The so obtained somehow clipped version of MGEs defines the well-known starting point for a stable semi-discretized version of the Maxwell's equations using FIT, see e. g. [SW01; CW02b], reading

$$\widetilde{S}\widehat{\widetilde{\mathbf{d}}} = \widehat{\widetilde{\mathbf{q}}}, \tag{4.31a}$$
$$S\widehat{\widetilde{\mathbf{b}}} = 0, \tag{4.31b}$$
$$C\widehat{\mathbf{e}} = -\frac{\partial}{\partial t}\widehat{\widetilde{\mathbf{b}}}, \tag{4.31c}$$
$$\widetilde{C}\widehat{\mathbf{h}} = \widehat{\widetilde{\mathbf{j}}} + \frac{\partial}{\partial t}\widehat{\widetilde{\mathbf{d}}}. \tag{4.31d}$$

The FIT builds upon the cell complexes' duality leading to a few important properties of the discrete operators which are crucial for the stability of numerical methods, cf. [Tho97]. Due to the cell complexes duality and the one-to-one relation, we obtain the following properties for the newly introduced discrete operators

$$\widetilde{G}_{ij} = -S^{\top}, \tag{4.32a}$$
$$\widetilde{S}_{ij} = -G^{\top}, \tag{4.32b}$$
$$\widetilde{C}_{ij} = C^{\top}. \tag{4.32c}$$

These properties, as well as (4.29) were shown and used for instance in [Wei77b; Wei84; Hah92; Cle98; Sch11]. Further, they allow for a semi-discrete version of the current continuity equation (4.3) by left-multiplying (4.31d) with $\widetilde{S}$ and inserting (4.31a) yielding

$$\widetilde{S}\widehat{\widetilde{\mathbf{j}}} + \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\widetilde{\mathbf{q}}} = 0.$$

Hence, charges are conserved by the spatially discretized Maxwell's equations. Concerning the implementation of these equations, consider the following remark.

Remark 4.18
The equations (4.28) are represented by ordinary vectors and matrices when fixing the index sets' order. Without loss of generality, let the p-cells of the primary cell complex $\mathcal{G}$ be labeled so that

$$\mathcal{N}_{\mathcal{P}} = \{1,\dots,|\mathcal{P}|\}, \quad \mathcal{N}_{\mathcal{L}} = \{1,\dots,|\mathcal{L}|\}, \quad \mathcal{N}_{\mathcal{F}} = \in \{1,\dots,|\mathcal{F}|\} \quad \textit{and} \quad \mathcal{N}_{\mathcal{V}} = \{1,\dots,|\mathcal{V}|\}.$$

Analogously, the dual cell complexes' labels shall be given by

$$\mathcal{N}_{\widetilde{\mathcal{P}}} = \{1,\dots,|\widetilde{\mathcal{P}}|\}, \quad \mathcal{N}_{\widetilde{\mathcal{L}}} = \{1,\dots,|\widetilde{\mathcal{L}}|\}, \quad \mathcal{N}_{\widetilde{\mathcal{F}}} = \in \{1,\dots,|\widetilde{\mathcal{F}}|\} \quad \textit{and} \quad \mathcal{N}_{\widetilde{\mathcal{V}}} = \{1,\dots,|\widetilde{\mathcal{V}}|\}$$

without violating the primal-dual counterpart relation for the same labeling as of Definition 4.11. Note that in this case, the dual p-cells of the boundary are labeled with indices grater then $|\mathcal{P}|, |\mathcal{L}|, |\mathcal{F}|$ *and* $|\mathcal{V}|$*, respectively for* $p \in \{0,1,2,3\}$*. Then, the discrete operators* $G, C, S, \widetilde{G}, \widetilde{C}$ *and* $\widetilde{S}$ *represent ordinary matrices. Further,* $\bar{\mathbf{e}}, \hat{\bar{\mathbf{b}}}, \hat{\mathbf{h}}, \hat{\bar{\mathbf{d}}}, \hat{\bar{\mathbf{j}}}$ *and* $\hat{\bar{\mathbf{q}}}$ *are ordinary vector functions.*

The reason why there are no fixed index orders yet is, that there are still boundary conditions to be projected into the discrete operators. Doing so changes the degrees of freedom which we address in the forthcoming sections.

Example 4.10 (continued)
We define integrals quantities on the primal or dual geometrical structures as follows. Quantities, which are related by the same equation, are assigned to p-cells of the same cell complex. Thus, we end up with

$$\begin{aligned}
\bar{\mathbf{e}}_{w,i_x,i_y,i_z}(t) &:= \int_{L(w,i_x,i_y,i_z)} \mathbf{E}(l,t)\cdot \mathrm{d}l && \forall (w,i_x,i_y,i_z) \in \mathcal{N}_{\mathcal{L}},\\
\hat{\bar{\mathbf{b}}}_{w,i_x,i_y,i_z}(t) &:= \iint_{F(w,i_x,i_y,i_z)} \mathbf{B}(s,t)\cdot \mathrm{d}s && \forall (w,i_x,i_y,i_z) \in \mathcal{N}_{\mathcal{F}},\\
\hat{\mathbf{h}}_{w,i_x,i_y,i_z}(t) &:= \int_{\widetilde{L}(w,i_x,i_y,i_z)} \mathbf{H}(l,t)\cdot \mathrm{d}l && \forall (w,i_x,i_y,i_z) \in \mathcal{N}_{\mathcal{F}},\\
\hat{\bar{\mathbf{d}}}_{w,i_x,i_y,i_z}(t) &:= \iint_{\widetilde{F}(w,i_x,i_y,i_z)} \mathbf{D}(s,t)\cdot \mathrm{d}s && \forall (w,i_x,i_y,i_z) \in \mathcal{N}_{\mathcal{L}},\\
\hat{\bar{\mathbf{j}}}_{w,i_x,i_y,i_z}(t) &:= \iint_{\widetilde{F}(w,i_x,i_y,i_z)} \mathbf{J}(s,t)\cdot \mathrm{d}s && \forall (w,i_x,i_y,i_z) \in \mathcal{N}_{\mathcal{L}},\\
\hat{\bar{\mathbf{q}}}_{i_x,i_y,i_z}(t) &:= \iiint_{\widetilde{V}(i_x,i_y,i_z)} \boldsymbol{\rho}(r,t)\,\mathrm{d}r && \forall (i_x,i_y,i_z) \in \mathcal{N}_{\mathcal{P}}.
\end{aligned}$$

Concerning the implementation issue of Remark 4.18, a proper vector of electric mesh voltages $\bar{\mathbf{e}}$ *can be obtained by fixing the order of* $\mathcal{N}_{\mathcal{L}}$ *by the following bijective mapping*

$$\begin{aligned}
m_{\mathcal{L}} &: \mathcal{N}_{\mathcal{L}} \to \{1,\dots,|\mathcal{N}_{\mathcal{L}}|\};\\
(w,i_x,i_y,i_z) &\mapsto \begin{cases}
i_x + (N_x-1)(i_y-1) + (N_x-1)N_y(i_z-1) & \text{if } w = \mathrm{x},\\
i_x + N_x(i_y-1) + N_x(N_y-1)(i_z-1) + |\mathcal{N}_{\mathcal{L}_x}| & \text{if } w = \mathrm{y},\\
i_x + N_x(i_y-1) + N_xN_y(i_z-1) + |\mathcal{N}_{\mathcal{L}_x}| + \left|\mathcal{N}_{\mathcal{L}_y}\right| & \text{if } w = \mathrm{z}.
\end{cases}
\end{aligned}$$

Hereby it is $|\mathcal{N}_{\mathcal{L}_x}| = (N_x-1)N_yN_z$ *and* $\left|\mathcal{N}_{\mathcal{L}_y}\right| = N_x(N_z-1)N_z$*. Then, the electric mesh voltages vector function is given by* $\bar{\mathbf{e}} := (\bar{\mathbf{e}}_{m_{\mathcal{L}}^{-1}(k)})_{k\in\{1,\dots,|\mathcal{N}_{\mathcal{L}}|\}}$*. The same mapping can be applied to the electric mesh fluxes* $\hat{\bar{\mathbf{d}}}$ *for the subset* $\mathcal{N}_{\mathcal{L}} \subset \mathcal{N}_{\widetilde{\mathcal{F}}}$*. Then, dual facets have their primal counterpart's index. Note that the dual boundary p-cells are not considered, but their role is eliminated after incorporating the boundary conditions in the forthcoming.*

The rectilinear mesh structure of both, the primal and dual cell complex, is advantageous for identifying incident $(p-1)$*-cells. For instance, Gauss's law for magnetism* (4.26b) *translates to*

$$-\hat{\bar{\mathbf{b}}}_{(\mathrm{x},i_x,i_y,i_z)} - \hat{\bar{\mathbf{b}}}_{(\mathrm{y},i_x,i_y,i_z)} - \hat{\bar{\mathbf{b}}}_{(\mathrm{z},i_x,i_y,i_z)} + \hat{\bar{\mathbf{b}}}_{(\mathrm{x},i_x+1,i_y,i_z)} + \hat{\bar{\mathbf{b}}}_{(\mathrm{y},i_x,i_y+1,i_z)} + \hat{\bar{\mathbf{b}}}_{(\mathrm{z},i_x,i_y,i_z+1)} = 0,$$

for $(i_x, i_y, i_z) \in \mathcal{N}_\mathcal{V}$*, whereby the orientation conventions are already incorporated, see Figure 4.14.*

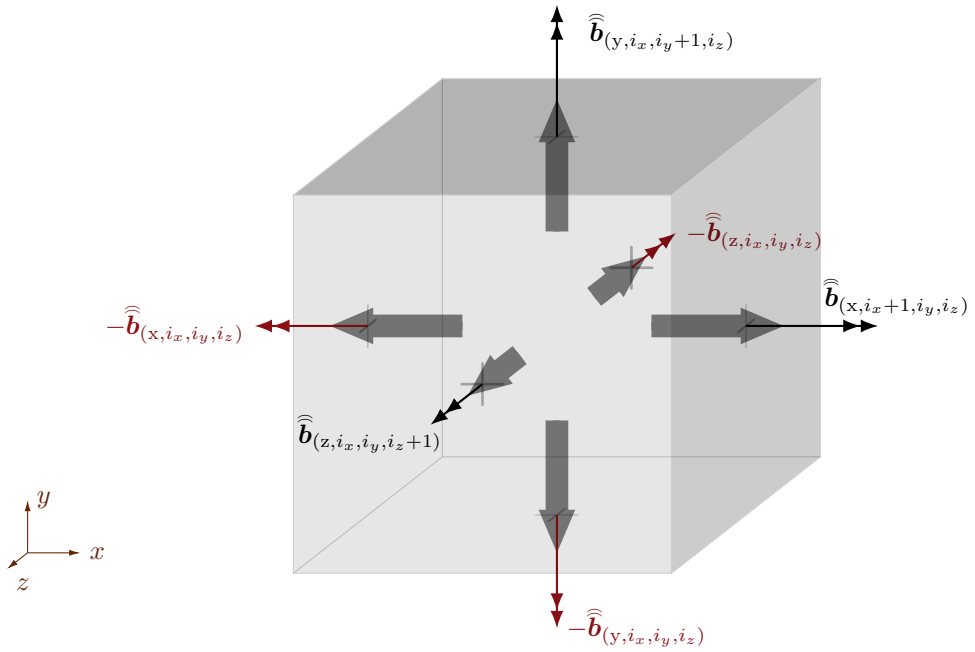


Figure 4.14: FIT visualisation of Gauss's law for magnetism. Red colored quantities are those which change sign due to orientation mismatch the according window function. The divergence for that volume's boundary follows the highlighted direction, i. e. outwards

Similar, Maxwell-Faraday's law (4.26c) reads, for $(w, i_x, i_y, i_z) \in \mathcal{N}_\mathcal{F}$*,*

- *if* $w = \mathrm{x}$*:*

$$\widehat{\mathbf{e}}_{(\mathrm{y},i_x,i_y,i_z)} + \widehat{\mathbf{e}}_{(\mathrm{z},i_x,i_y+1,i_z)} - \widehat{\mathbf{e}}_{(\mathrm{y},i_x,i_y,i_z+1)} - \widehat{\mathbf{e}}_{(\mathrm{z},i_x,i_y,i_z)} = -\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\widehat{\mathbf{b}}}_{(\mathrm{x},i_x,i_y,i_z)},$$

- *if* $w = \mathrm{y}$*:*

$$\widehat{\mathbf{e}}_{(\mathrm{z},i_x,i_y,i_z)} + \widehat{\mathbf{e}}_{(\mathrm{x},i_x,i_y,i_z+1)} - \widehat{\mathbf{e}}_{(\mathrm{z},i_x+1,i_y,i_z)} - \widehat{\mathbf{e}}_{(\mathrm{x},i_x,i_y,i_z)} = -\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\widehat{\mathbf{b}}}_{(\mathrm{y},i_x,i_y,i_z)},$$

- *if* $w = \mathrm{z}$*:*

$$\widehat{\mathbf{e}}_{(\mathrm{x},i_x,i_y,i_z)} + \widehat{\mathbf{e}}_{(\mathrm{y},i_x+1,i_y,i_z)} - \widehat{\mathbf{e}}_{(\mathrm{x},i_x,i_y+1,i_z)} - \widehat{\mathbf{e}}_{(\mathrm{y},i_x,i_y,i_z)} = -\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\widehat{\mathbf{b}}}_{(\mathrm{z},i_x,i_y,i_z)}$$

as shown in Figure 4.15 for the latter case. Gauss's law and Maxwell-Ampère's law are obtained in a similar fashion by switching to the dual cell complex. Note that these laws are given in the labeling domain and, again, using the index mapping requires to treat the dual boundary cells with precaution.

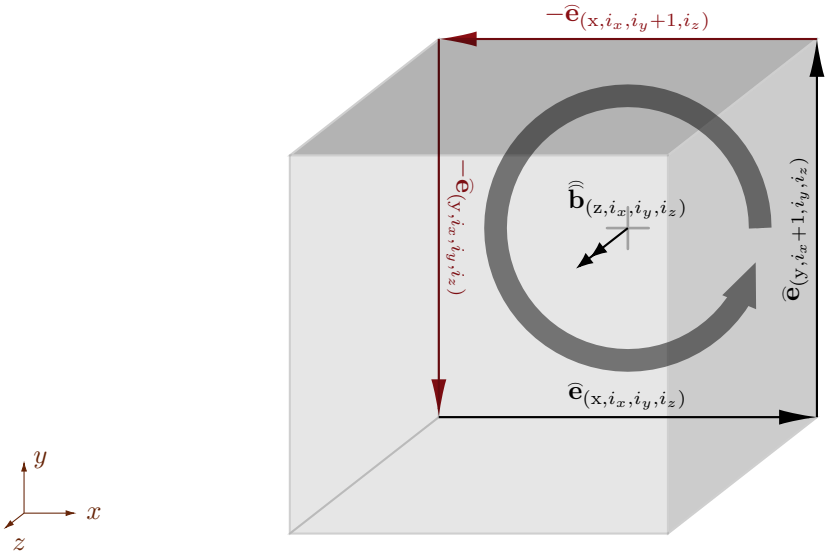


Figure 4.15: FIT visualisation of Maxwell-Faraday's law. Red colored quantities are those which change sign due to orientation by the according window function. The curl for $\widehat{\widehat{\mathbf{b}}}$ follows the highlighted direction (Ampère's right-hand grip rule)

4.5.5 Constitutive Grid Relations

In order to complete MGEs, we require spatially discretized constitutive relations which are called constitutive grid relations (CGRs). These CGRs establish a coupling between the primal and dual grid by relating their mesh quantities, e. g. they relate $\widehat{\mathbf{e}}$ with $\widehat{\widehat{\mathbf{d}}}$ or $\widehat{\mathbf{h}}$ and $\widehat{\widehat{\mathbf{b}}}$. Due to the staggered grid design, primal and dual mesh quantities are already associated with each other by their underlying geometrical objects' intersections. Recall that the primal-dual counterpart role only applies to the dual p-cells that are not subset to the boundary, in other words those which share the same primal counterpart's label. However, these associations involve the coupling of quantities on different geometrical structures such as the line integrals $\widehat{\mathbf{e}}$ with their associated surface integrals $\widehat{\widehat{\mathbf{d}}}$. The coupling is archived by the so-called material matrices which are constructed by making use of average determination and metrical information. This is exactly where the p-cells geometry produces an error, see e. g. [Cle98; CW01b; Bau12], contrary to validity of MEs in the scope of each p-cell as of Remark 4.17.

In order to reason with averaging we require the following essential assumption.

Assumption 4.19
On each primal volume $V(i) \in \mathcal{V}$ the materials are assumed to be homogeneous.

Under the prerequisite of Assumption 4.19, we can exploit the constitutive relations (4.6) in order to derive approximations which serve as the discretized version of them as follows. We start by defining

$$h := \max_{L \in \mathcal{L}} \int_L 1 \, \mathrm{d}l.$$

Let $|E_i|$ be the sample magnitude of the electric field strength $\mathbf{E}$ along $L(i)$ at the intersection with $\widetilde{F}(i)$, for $i \in \mathcal{N}_{\mathcal{L}}$. Then, we obtain the approximations

$$\frac{\widehat{\mathbf{d}}_i}{\widehat{\mathbf{e}}_i} = \frac{\iint_{\widetilde{F}(i)} \varepsilon(s)\mathbf{E}(s,t)\cdot \mathrm{d}s}{\int_{L(i)} \mathbf{E}(l,t)\cdot \mathrm{d}l} = \frac{\bar{\varepsilon}_i \iint_{\widetilde{F}(i)} |E_i|\,\mathrm{d}s}{\int_{L(i)} |E_i|\,\mathrm{d}l} + \mathcal{O}(h^{k+1}) \approx \bar{\varepsilon}_i \frac{\iint_{\widetilde{F}(i)} 1\,\mathrm{d}s}{\int_{L(i)} 1\,\mathrm{d}l},$$
$$\frac{\widehat{\mathbf{j}}_i}{\widehat{\mathbf{e}}_i} = \frac{\iint_{\widetilde{F}(i)} \sigma(s)\mathbf{E}(s,t)\cdot \mathrm{d}s}{\int_{L(i)} \mathbf{E}(l,t)\cdot \mathrm{d}l} = \frac{\bar{\sigma}_i \iint_{\widetilde{F}(i)} |E_i|\,\mathrm{d}s}{\int_{L(i)} |E_i|\,\mathrm{d}l} + \mathcal{O}(h^{k+1}) \approx \bar{\sigma}_i \frac{\iint_{\widetilde{F}(i)} 1\,\mathrm{d}s}{\int_{L(i)} 1\,\mathrm{d}l}$$

whereby $\bar{\varepsilon}_i := \frac{\iint_{\widetilde{F}(i)} \varepsilon(s)\,\mathrm{d}s}{\iint_{\widetilde{F}(i)} 1\,\mathrm{d}s}$ and $\bar{\sigma}_i := \frac{\iint_{\widetilde{F}(i)} \sigma(s)\,\mathrm{d}s}{\iint_{\widetilde{F}(i)} \mathrm{d}s}$ are the average permittivity and conductivity on $\widetilde{F}(i)$, respectively, and is the convergence order of the FIT approximation. The convergence order depends on cell complexes' structure and materials, for instance k may reach order 2 in case of equidistant meshes spanned over a domain with a homogeneous material where the sample points are in located at the 1 and 2-cell's center - for a deeper insight see further [Hah92; Krü00]. The next coupling approach differs slightly since we now consider $|B_i|$ which is the sample magnitude of the magnetic flux density $\mathbf{B}$ along the normal field of $F(i)$ at the intersection with $\widetilde{L}(i)$, for $i \in \mathcal{N}_{\mathcal{F}}$. In a similar fashion, we obtain

$$\frac{\widehat{\mathbf{h}}_i}{\widehat{\mathbf{b}}_i} = \frac{\int_{\widetilde{L}(i)} \nu(s)\mathbf{B}(l,t)\cdot \mathrm{d}l}{\iint_{F(i)} \mathbf{B}(s,t)\cdot \mathrm{d}s} = \frac{\bar{\nu}_i \int_{\widetilde{L}(i)} |B_i|\,\mathrm{d}l}{\iint_{F(i)} |B_i|\,\mathrm{d}s} + \mathcal{O}(h^{k+2}) \approx \bar{\nu}_i \frac{\int_{\widetilde{L}(i)} 1\,\mathrm{d}l}{\iint_{F(i)} 1\,\mathrm{d}s}$$

with average inverse permeability $\bar{\nu}_i := \frac{\int_{\widetilde{L}(i)} \nu(s)\,\mathrm{d}s}{\int_{\widetilde{L}(i)} \mathrm{d}s}$, see e. g. [Wei96]. Note that we made use of Assumption 4.15 which allows us to use the sample magnitudes straight forward for the dual counterparts.

All metric information are collected in the diagonal matrices

$$D_{\mathcal{L}} := \operatorname{diag}\Big(\int_{L(i)} 1\,\mathrm{d}l\Big)_{i\in\mathcal{N}_{\mathcal{L}}}, \quad D_{\mathcal{F}} := \operatorname{diag}\Big(\iint_{F(i)} 1\,\mathrm{d}s\Big)_{i\in\mathcal{N}_{\mathcal{F}}},$$
$$D_{\widetilde{\mathcal{L}}} := \operatorname{diag}\Big(\int_{\widetilde{L}(i)} 1\,\mathrm{d}l\Big)_{i\in\mathcal{N}_{\mathcal{F}}}, \quad D_{\widetilde{\mathcal{F}}} := \operatorname{diag}\Big(\iint_{\widetilde{F}(i)} 1\,\mathrm{d}s\Big)_{i\in\mathcal{N}_{\mathcal{L}}}.$$

Since the materials of consideration are isotropic by Assumption 4.3 and the cell complexes are orthogonal to each other, again by Assumption 4.15, the average material properties are scalars and can be collected in the diagonal matrices

$$D_\varepsilon := \operatorname{diag}(\bar{\varepsilon}_i)_{i\in\mathcal{N}_{\mathcal{L}}}, \quad D_\nu := \operatorname{diag}(\bar{\nu}_i)_{i\in\mathcal{N}_{\mathcal{F}}} \quad \text{and} \quad D_\sigma := \operatorname{diag}(\bar{\sigma}_i)_{i\in\mathcal{N}_{\mathcal{L}}}.$$

Together we introduce the material matrices

$$M_\varepsilon := D_{\widetilde{\mathcal{F}}} D_\varepsilon D_{\mathcal{L}}^{-1}, \qquad M_\nu := D_{\mathcal{F}} D_\nu D_{\widetilde{\mathcal{L}}}^{-1} \qquad \text{and} \qquad M_\sigma := D_{\widetilde{\mathcal{F}}} D_\sigma D_{\mathcal{L}}^{-1}.$$

Note that for nonorthogonal cell complexes the material matrices would be of band matrix structure but no longer diagonal [SW98b; SW98a; SW99; Sch99]. An approach for anisotropic materials is given in [Krü00; Gut98]. Further it shall be mentioned that the inverses are well

defined since especially the 1- and 2-cells are not allowed to be null spaces.
With these shorthands, the just introduced approximative model yields the *constitutive grid relations*

$$\widehat{\widehat{\mathbf{d}}} = M_\varepsilon \widehat{\mathbf{e}}, \tag{4.33a}$$
$$\widehat{\widehat{\mathbf{j}}} = M_\sigma \widehat{\mathbf{e}}, \tag{4.33b}$$
$$\widehat{\mathbf{h}} = M_\nu \widehat{\widehat{\mathbf{b}}}. \tag{4.33c}$$

A typical challenge of bringing both cell complexes' quantities together is to improve the geometric approximation of actual material shapes (staircase approximation problem when using Cartesian meshes). Possible improvements are for instance subgridding techniques or the use of different subvolume techniques making use of triangular or tetrahedral partitioning, see e. g. [Wei79; TW96] or the *conformal finite integration technique* [CW02a] for curved material surfaces.

Example 4.10 (continued)
Due to the rectilinear structure of the meshes, we can easily obtain all the integral areas from the primal coordinates. With further Assumption 4.19 to hold, each $\widehat{\widehat{\mathbf{d}}}_i$*, for* $i \in \mathcal{N}_\mathcal{L}$*, depends only on up to four values of* ε*, one for each incident volume quadrant, denoted with* $\varepsilon_\mathrm{I}, \varepsilon_\mathrm{II}, \varepsilon_\mathrm{III}$ *and* ε_IV*. The same holds for* $\widehat{\widehat{\mathbf{j}}}_i$ *with* $\sigma_\mathrm{I}, \sigma_\mathrm{II}, \sigma_\mathrm{III}$ *and* σ_IV*. To be more precise, consider (4.33a) and let* $i = (w, i_x, i_y, i_z) \in \mathcal{N}_\mathcal{L}$*. Given that* $w = \mathrm{z}$ *it is*

$$\varepsilon_\mathrm{I} = \varepsilon\left(\begin{pmatrix} \widetilde{x}_{i_x+1} \\ \widetilde{y}_{i_y+1} \\ \widetilde{z}_{i_z+1} \end{pmatrix}\right), \quad \varepsilon_\mathrm{II} = \varepsilon\left(\begin{pmatrix} \widetilde{x}_{i_x} \\ \widetilde{y}_{i_y+1} \\ \widetilde{z}_{i_z+1} \end{pmatrix}\right), \quad \varepsilon_\mathrm{III} = \varepsilon\left(\begin{pmatrix} \widetilde{x}_{i_x} \\ \widetilde{y}_{i_y} \\ \widetilde{z}_{i_z+1} \end{pmatrix}\right), \quad \varepsilon_\mathrm{IV} = \varepsilon\left(\begin{pmatrix} \widetilde{x}_{i_x+1} \\ \widetilde{y}_{i_y} \\ \widetilde{z}_{i_z+1} \end{pmatrix}\right)$$

where the indices are in bound; otherwise the according terms do not exist for instance at the boundary. Then, the individual equation line for i *reads*

$$\widehat{\widehat{\mathbf{d}}}_i = \Big[\frac{(x_{i_x+1} - x_{i_x})(y_{i_y+1} - y_{i_y})}{4}\varepsilon_\mathrm{I} + \frac{(x_{i_x} - x_{i_x-1})(y_{i_y+1} - y_{i_y})}{4}\varepsilon_\mathrm{II} + \frac{(x_{i_x} - x_{i_x-1})(y_{i_y} - y_{i_y-1})}{4}\varepsilon_\mathrm{III} + \frac{(x_{i_x+1} - x_{i_x})(y_{i_y} - y_{i_y-1})}{4}\varepsilon_\mathrm{IV}\Big]\frac{\widehat{\mathbf{e}}_i}{(z_{i_z+1} - z_{i_z})},$$

see Figure 4.16 for a visual interpretation. The expression in square brackets is the average permittivity $\bar{\varepsilon}_i$ *multiplied by the area of* $F(i)$ *which is canceled out. In an entire analogous fashion, we obtain the equations for* $w \in \{\mathrm{x}, \mathrm{y}\}$ *and also (4.33b).*
In case of (4.33c), $\widehat{\mathbf{h}}_i$*, for* $i = (w, i_x, i_y, i_z) \in \mathcal{N}_{\widetilde{\mathcal{L}}}$*, depends only on up to two values of the reluctivity, that is* ν_I *and* ν_II*. For the case* $w = \mathrm{y}$ *we have for instance*

$$\nu_\mathrm{I} = \nu\left(\begin{pmatrix} \widetilde{x}_{i_x+1} \\ \widetilde{y}_{i_y+1} \\ \widetilde{z}_{i_z+1} \end{pmatrix}\right), \quad \nu_\mathrm{II} = \nu\left(\begin{pmatrix} \widetilde{x}_{i_x+1} \\ \widetilde{y}_{i_y} \\ \widetilde{z}_{i_z+1} \end{pmatrix}\right),$$

and the equation line for i *reads*

$$\widehat{\mathbf{h}}_i = \Big[\frac{(y_{i_y+1} - y_{i_y})}{2}\nu_\mathrm{I} + \frac{(y_{i_y} - y_{i_y-1})}{2}\nu_\mathrm{II}\Big]\frac{\widehat{\widehat{\mathbf{b}}}_i}{(x_{i_x+1} - x_{i_x})(z_{i_z+1} - z_{i_z})},$$

see Figure 4.17, whereby the expression in square brackets is the average reluctivity $\bar{\nu}_i$ *multiplied by the length of* $\widetilde{L}(i)$ *which is canceled out in the overall expression.*

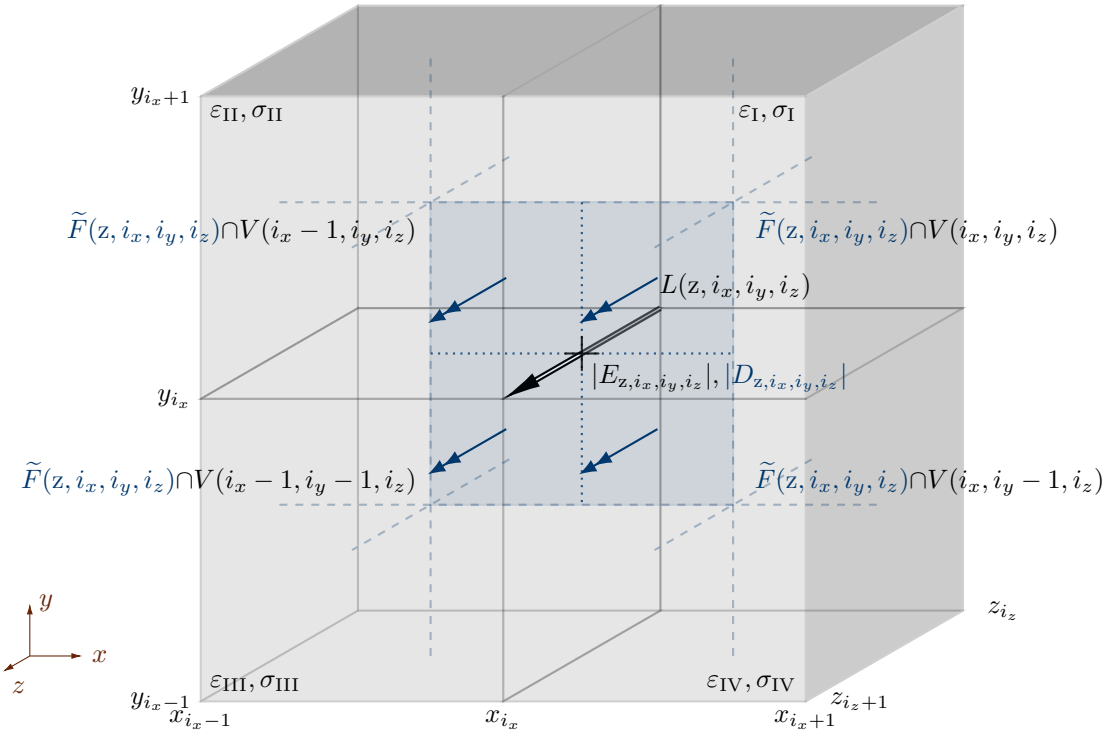


Figure 4.16: FIT visualization of constitutive grid relation linking primal link and dual facet, representative for (4.33a) and (4.33b).

4.5.6 Boundary Grid Conditions

As mentioned in Section 4.5.4, the MGEs do not consider boundary conditions yet. In fact, MGEs together with the CGRs still describe an under-determined system of equations. In this treatise we focus on PEC boundaries and want to incorporate the spatially discretized version of these boundary conditions (4.13) into MGEs

$$\widehat{\mathbf{e}}_i = 0 \qquad \forall i \in \mathcal{N}_{\mathcal{L}} \ : \ L(i) \subset \partial\Omega, \tag{4.34a}$$
$$\widehat{\widetilde{\mathbf{d}}}_i = M_\varepsilon \widehat{\mathbf{e}}_i = 0 \qquad \forall i \in \mathcal{N}_{\mathcal{L}} \ : \ L(i) \subset \partial\Omega, \tag{4.34b}$$
$$\widehat{\widetilde{\mathbf{j}}}_i = M_\sigma \widehat{\mathbf{e}}_i = 0 \qquad \forall i \in \mathcal{N}_{\mathcal{L}} \ : \ L(i) \subset \partial\Omega, \tag{4.34c}$$
$$\widehat{\widetilde{\mathbf{b}}}_i = 0 \qquad \forall i \in \mathcal{N}_{\mathcal{F}} \ : \ F(i) \subset \partial\Omega, \tag{4.34d}$$
$$\widehat{\mathbf{h}}_i = M_\nu \widehat{\widetilde{\mathbf{b}}}_i = 0 \qquad \forall i \in \mathcal{N}_{\mathcal{F}} \ : \ F(i) \subset \partial\Omega. \tag{4.34e}$$

This motivates the introduction of reduced index sets

$$\mathcal{N}_{\mathcal{P}}^{\mathrm{bound}} := \{ i \in \mathcal{N}_{\mathcal{P}} \mid P(i) \subset \partial\Omega \}, \qquad \mathcal{N}_{\mathcal{P}}^{\mathrm{int}} := \mathcal{N}_{\mathcal{P}} \backslash \mathcal{N}_{\mathcal{P}}^{\mathrm{bound}},$$
$$\mathcal{N}_{\mathcal{L}}^{\mathrm{bound}} := \{ i \in \mathcal{N}_{\mathcal{L}} \mid L(i) \subset \partial\Omega \}, \qquad \mathcal{N}_{\mathcal{L}}^{\mathrm{int}} := \mathcal{N}_{\mathcal{L}} \backslash \mathcal{N}_{\mathcal{L}}^{\mathrm{bound}},$$
$$\mathcal{N}_{\mathcal{F}}^{\mathrm{bound}} := \{ i \in \mathcal{N}_{\mathcal{F}} \mid F(i) \subset \partial\Omega \}, \qquad \mathcal{N}_{\mathcal{F}}^{\mathrm{int}} := \mathcal{N}_{\mathcal{F}} \backslash \mathcal{N}_{\mathcal{F}}^{\mathrm{bound}}$$

for the forthcoming.

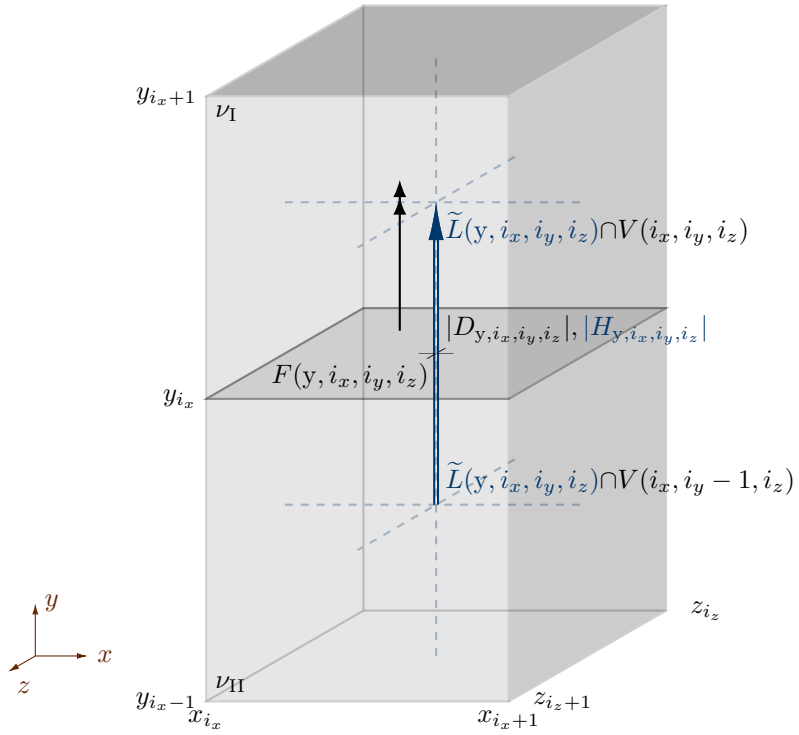


Figure 4.17: FIT visualization of constitutive grid relation linking primal facet and dual link, representative for (4.33c).

4.5.7 Maxwell's Grid Equations in Potential Formulation

Following the potential approach in the continuous case, we analogously introduce the *discrete magnetic vector potential* $\widehat{\mathbf{a}} : \mathcal{I} \to \mathbb{R}^{\mathcal{N}_\mathcal{L}}$ so that

$$\widehat{\widehat{\mathbf{b}}} = C\widehat{\mathbf{a}} \tag{4.35}$$

and we obtain with the same argumentation as in the continuous case

$$\widehat{\mathbf{e}} = -G\boldsymbol{\phi} - \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{a}} \tag{4.36}$$

where $\boldsymbol{\phi} : \mathcal{I} \to \mathbb{R}^{\mathcal{N}_\mathcal{P}}$ is the *discrete electric scalar potential*. Substituting (4.35) and (4.36) into (4.31) and exploiting (4.33) yields a semi-discrete Gauss's law and Maxwell-Ampère's law in discrete potential quantities

$$-\widetilde{S}M_\varepsilon\big(G\boldsymbol{\phi} + \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{a}}\big) = \widehat{\widehat{\mathbf{q}}}, \tag{4.37a}$$

$$\widetilde{C}M_\nu C\widehat{\mathbf{a}} = -\frac{\mathrm{d}}{\mathrm{d}t}\Big[M_\varepsilon\big(G\boldsymbol{\phi} + \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{a}}\big)\Big] - M_\sigma\big(G\boldsymbol{\phi} + \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{a}}\big). \tag{4.37b}$$

The remaining equations (4.31b) and (4.31c) are implicitly fulfilled by the discrete operators properties (4.29) and (4.32).

In terms of discrete potentials, the total current flow (4.30) through a dual surface $\widetilde{F}(i)$, for $i \in \mathcal{N}_{\mathcal{L}}$, reads

$$\Big(-\frac{\mathrm{d}}{\mathrm{d}t}\big[M_\varepsilon\big(G\boldsymbol{\phi} + \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{a}}\big)\big] - M_\sigma\big(G\boldsymbol{\phi} + \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{a}}\big)\Big)_i = \big(\widetilde{C}M_\nu C\widehat{\mathbf{a}}\big)_i. \tag{4.38}$$

In analogy to Maxwell's house for the continuous quantities, for the spatial discretized quantities and operators the relations are given in Figure 4.18.

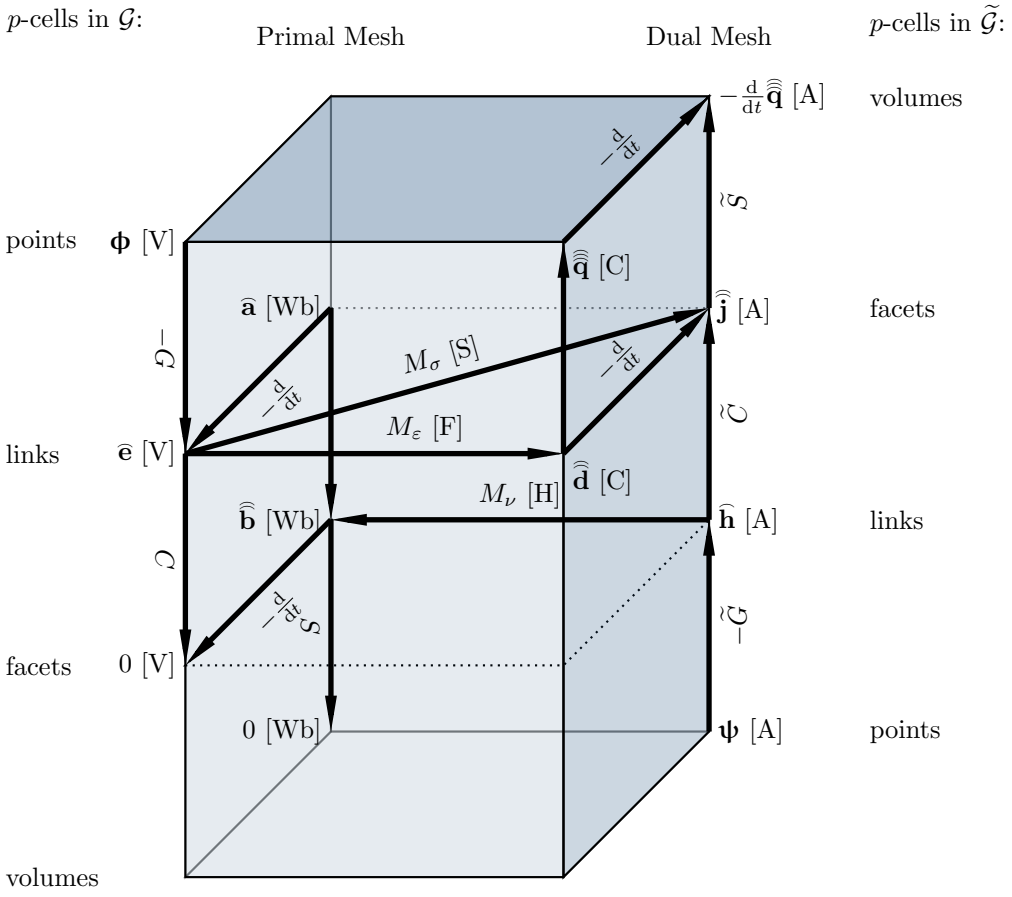


Figure 4.18: Maxwell's house or Tonti's diagram for spatially discretized electromagnetic quantities.

Boundary Conditions and Excitation Assuming PEC boundaries, that is Assumption 4.6 for the $\mathbf{A} - \boldsymbol{\varphi}$ formulation, motivates the following discrete realizations:

(i) The discrete vector potentials tangential to the boundary vanish, i. e.

$$\widehat{\mathbf{a}}_i = 0, \quad \text{for } i \in \mathcal{N}_{\mathcal{L}}^{\text{bound}}. \tag{4.39}$$

(ii) The scalar potentials gradient tangential to the boundary vanishes, i. e.

$$(G_{ij})_{i\in\mathcal{N}_{\mathcal{L}}^{\text{bound}},j\in\mathcal{N}_{\mathcal{P}}}\boldsymbol{\Phi}=0. \tag{4.40}$$

The latter condition (4.40) motivates to set the boundary potentials to an almost everywhere spatially constant function possibly varying in time

$$\boldsymbol{\Phi}_i=\phi_{\Gamma,i},\qquad \text{for } i\in\mathcal{N}_{\mathcal{P}}^{\text{bound}} \tag{4.41}$$

whereby $\phi_\Gamma:\mathcal{I}\to\mathbb{R}^{\mathcal{N}_{\mathcal{P}}^{\text{bound}}}$ is a given function serving as some excitation. Given these boundary conditionswe eliminate all entries of the discrete operators, material matrices and integral quantities that vanish. For this we introduce the following reduced discrete operators, material matrices and integral quantities

$$\begin{aligned}
G^{\text{red}}&:=(G_{,ij})_{i\in\mathcal{N}_{\mathcal{L}}^{\text{int}},j\in\mathcal{N}_{\mathcal{P}}^{\text{int}}}, & G_\Gamma&:=(G_{\text{exa},ij})_{i\in\mathcal{N}_{\mathcal{L}}^{\text{int}},j\in\mathcal{N}_{\mathcal{P}}^{\text{bound}}},\\
\widetilde{S}^{\text{red}}&:=(\widetilde{S}_{\text{exa},ij})_{i\in\mathcal{N}_{\mathcal{P}}^{\text{int}},j\in\mathcal{N}_{\mathcal{L}}^{\text{int}}}, & \widetilde{S}_\Gamma&:=(\widetilde{S}_{\text{exa},ij})_{i\in\mathcal{N}_{\mathcal{P}}^{\text{bound}},j\in\mathcal{N}_{\mathcal{L}}^{\text{int}}},\\
C^{\text{red}}&:=(C_{,ij})_{i\in\mathcal{N}_{\mathcal{F}}^{\text{int}},j\in\mathcal{N}_{\mathcal{L}}^{\text{int}}},\\
\widetilde{C}^{\text{red}}&:=(\widetilde{C}_{\text{exa},ij})_{i\in\mathcal{N}_{\mathcal{L}}^{\text{int}},j\in\mathcal{N}_{\mathcal{F}}^{\text{int}}},\\
M_\varepsilon^{\text{red}}&:=(M_{\varepsilon,ij})_{i\in\mathcal{N}_{\mathcal{L}}^{\text{int}},j\in\mathcal{N}_{\mathcal{L}}^{\text{int}}}, & \widehat{\widehat{\mathbf{d}}}^{\text{red}}&:=(\widehat{\widehat{\mathbf{d}}}_i)_{i\in\mathcal{N}_{\mathcal{L}}^{\text{int}}},\\
M_\sigma^{\text{red}}&:=(M_{\sigma,ij})_{i\in\mathcal{N}_{\mathcal{L}}^{\text{int}},j\in\mathcal{N}_{\mathcal{L}}^{\text{int}}}, & \widehat{\widehat{\mathbf{j}}}^{\text{red}}&:=(\widehat{\widehat{\mathbf{d}}}_i)_{i\in\mathcal{N}_{\mathcal{L}}^{\text{int}}},\\
M_\nu^{\text{red}}&:=(M_{\nu,ij})_{i\in\mathcal{N}_{\mathcal{F}}^{\text{int}},j\in\mathcal{N}_{\mathcal{F}}^{\text{int}}}, & \widehat{\mathbf{h}}^{\text{red}}&:=(\widehat{\widehat{\mathbf{d}}}_i)_{i\in\mathcal{N}_{\mathcal{F}}^{\text{int}}},\\
\widehat{\mathbf{a}}^{\text{red}}&:=(\widehat{\mathbf{a}}_i)_{i\in\mathcal{N}_{\mathcal{L}}^{\text{int}}}, & \widehat{\mathbf{e}}^{\text{red}}&:=(\widehat{\mathbf{e}}_i)_{i\in\mathcal{N}_{\mathcal{L}}^{\text{int}}},\\
\boldsymbol{\Phi}^{\text{red}}&:=(\boldsymbol{\Phi}_i)_{i\in\mathcal{N}_{\mathcal{P}}^{\text{int}}}, & \widehat{\widehat{\mathbf{b}}}^{\text{red}}&:=(\widehat{\widehat{\mathbf{b}}}_i)_{i\in\mathcal{N}_{\mathcal{F}}^{\text{int}}}.
\end{aligned}$$

After incorporating the discrete boundary conditions and excitation, that is (4.39) and (4.40), from (4.35) and (4.36) we obtain respectively

$$\widehat{\widehat{\mathbf{b}}}^{\text{red}}=C^{\text{red}}\widehat{\mathbf{a}}^{\text{red}},\qquad\qquad \widehat{\mathbf{e}}^{\text{red}}=-G^{\text{red}}\boldsymbol{\Phi}^{\text{red}}-G_\Gamma\phi_\Gamma(t)-\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{a}}^{\text{red}}$$

and the remaining constitutive grid equations read

$$\widehat{\widehat{\mathbf{d}}}^{\text{red}}=M_\varepsilon^{\text{red}}\widehat{\mathbf{e}}^{\text{red}},\qquad\qquad \widehat{\widehat{\mathbf{j}}}^{\text{red}}=M_\sigma^{\text{red}}\widehat{\mathbf{e}}^{\text{red}},\qquad\qquad \widehat{\mathbf{h}}^{\text{red}}=M_\nu^{\text{red}}\widehat{\widehat{\mathbf{b}}}^{\text{red}}.$$

Note that these essential equations comply with the boundary conditions for the standard formulation (4.34). Following an analogous procedure as for (4.37), we obtain the semi-discrete Gauss's law and Maxwell-Ampère's law in discrete potential quantities

$$-\widetilde{S}^{\text{red}}M_\varepsilon^{\text{red}}\big(G^{\text{red}}\boldsymbol{\Phi}^{\text{red}}+\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{a}}^{\text{red}}\big)=\widehat{\widehat{\mathbf{q}}}^{\text{red}}, \tag{4.42a}$$

$$\begin{aligned}
\widetilde{C}^{\text{red}}M_\nu^{\text{red}}C^{\text{red}}\widehat{\mathbf{a}}^{\text{red}}=&-\frac{\mathrm{d}}{\mathrm{d}t}\Big[M_\varepsilon^{\text{red}}\big(G^{\text{red}}\boldsymbol{\Phi}^{\text{red}}+G_\Gamma\phi_\Gamma(t)+\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{a}}^{\text{red}}\big)\Big]\\
&-M_\sigma^{\text{red}}\big(G^{\text{red}}\boldsymbol{\Phi}^{\text{red}}+G_\Gamma\phi_\Gamma(t)+\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{a}}^{\text{red}}\big).
\end{aligned} \tag{4.42b}$$

In terms of discrete potentials, the total current flow (4.30) through a dual surfaces $\widetilde{F}(i)$, for $i \in \mathcal{N}_{\mathcal{L}}$, reads, for instance

$$\left(\widetilde{C}^{\mathrm{red}} M_{\nu}^{\mathrm{red}} C^{\mathrm{red}} \widehat{\mathbf{a}}^{\mathrm{red}}\right)_i. \tag{4.43}$$

Gauge Fixing Similar as pointed out in Section 4.4.2, even with incorporated boundary conditions, the system (4.42) does not uniquely determine $\bar{\mathbf{a}}^{\mathrm{red}}$ and $\boldsymbol{\phi}^{\mathrm{red}}$. With focus on Lorenz gauge, we transfer the generalized gauge condition (4.20), for $\vartheta = 1$, to the semi-discrete version, see [Bau12],

$$M_{\varepsilon}^{\mathrm{red}} G^{\mathrm{red}} \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\phi}^{\mathrm{red}} + M_{\zeta}^{\mathrm{red}} G^{\mathrm{red}} M_{\xi}^{\mathrm{red}} \widetilde{S}^{\mathrm{red}} M_{\zeta}^{\mathrm{red}} \widehat{\mathbf{a}}^{\mathrm{red}} = 0, \tag{4.44}$$

whereby $M_{\zeta}^{\mathrm{red}} \in \mathbb{R}^{\mathcal{N}_{\mathcal{L}}^{\mathrm{int}} \times \mathcal{N}_{\mathcal{L}}^{\mathrm{int}}}$ and $M_{\xi}^{\mathrm{red}} \in \mathbb{R}^{\mathcal{N}_{\mathcal{P}}^{\mathrm{int}} \times \mathcal{N}_{\mathcal{P}}^{\mathrm{int}}}$ are artificial material matrices. The numerical results in Section 7 are produced by setting

$$\begin{aligned} M_{\zeta}^{\mathrm{red}} &:= \mathrm{diag}\left(M_{\zeta,i}\right)_{i \in \mathcal{N}_{\mathcal{P}}^{\mathrm{int}}} \;:\; M_{\zeta,i} = \begin{cases} M_{\sigma ii} & \text{if } M_{\sigma,ii} \neq 0 \\ 1 & \text{else} \end{cases}, \\ M_{\xi}^{\mathrm{red}} &:= \mathrm{diag}\left(M_{\xi,i}\right)_{i \in \mathcal{N}_{\mathcal{P}}^{\mathrm{int}}} \;:\; M_{\xi,i} = \frac{1}{(\widetilde{S} M_{\varepsilon} G)_{ii}}. \end{aligned}$$

MEs in Potential Formulation on Mesh with PEC Collecting all the essentials parts from above, that is (4.42b) and (4.44), a possible spatially discretized electromagnetic model for the $\mathbf{A} - \boldsymbol{\varphi}$ formulation (4.21), considered in [Bau12], is given by the following system of equations

$$\widetilde{S}^{\mathrm{red}} M_{\varepsilon}^{\mathrm{red}} G^{\mathrm{red}} \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\phi}^{\mathrm{red}} + \widetilde{S}^{\mathrm{red}} M_{\zeta}^{\mathrm{red}} G^{\mathrm{red}} M_{\xi}^{\mathrm{red}} \widetilde{S}^{\mathrm{red}} M_{\zeta}^{\mathrm{red}} \widehat{\mathbf{a}}^{\mathrm{red}} = 0, \tag{4.45a}$$

$$\begin{aligned} \widetilde{C}^{\mathrm{red}} M_{\nu}^{\mathrm{red}} C^{\mathrm{red}} \widehat{\mathbf{a}}^{\mathrm{red}} + \frac{\mathrm{d}}{\mathrm{d}t} \Big[M_{\varepsilon}^{\mathrm{red}} \big(G^{\mathrm{red}} \boldsymbol{\phi}^{\mathrm{red}} + G_{\Gamma} \phi_{\Gamma}(t) + \boldsymbol{\pi}^{\mathrm{red}} \big) \Big] \\ + M_{\sigma}^{\mathrm{red}} \big(G^{\mathrm{red}} \boldsymbol{\phi}^{\mathrm{red}} + G_{\Gamma} \phi_{\Gamma}(t) + \frac{\mathrm{d}}{\mathrm{d}t} \widehat{\mathbf{a}}^{\mathrm{red}} \big) = 0, \end{aligned} \tag{4.45b}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \widehat{\mathbf{a}}^{\mathrm{red}} - \boldsymbol{\pi}^{\mathrm{red}} = 0. \tag{4.45c}$$

where $\boldsymbol{\pi}^{\mathrm{red}} : \mathcal{I} \to \mathbb{R}^{\mathcal{N}_{\mathcal{L}}^{\mathrm{int}}}$ is the *discrete quasi-canonical momentum* introduced as an auxiliary function to avoid second order time differentiation of $\bar{\mathbf{a}}^{\mathrm{red}}$, in analogy to the continuous case (4.19).

Hereinafter, we denote with MGEs the system (4.45) which is actually one possible form of Maxwell's grid equations 4.26 expressed in terms of potentials, using Lorenz gauge and incorporated PEC boundary conditions with additional excitation. Note that $\widehat{\widehat{\mathbf{q}}}^{\mathrm{red}}$is now an output variable which can be obtained by (4.42a).

Underlying Assumptions and Properties The MGEs (4.45) represent reasonable model equations to solve the electromagnetic problem

- for certain materials Assumption 4.3;
- on finite three-dimensional domain Ω Assumption 4.5;
- surrounded by perfectly conducting material Assumption 4.6

which is spatially discretized using FIT with staggered computational meshes

- fulfilling each Assumption 4.8
- are orthogonal to each other Assumption 4.15
- where the primal one follows the material Assumption 4.19

Having chosen the domain's shape, computational meshes and boundary conditions carefully while using physically reasonable material, the following properties, which we collect in an assumption, can be proven, see for instance [Wei84; Hah92; Rie01].

Assumption 4.20 (properties of MGEs)

(i) *The permittivity and reluctivity matrices M_ε and M_ν are symmetric positive definite.*

(ii) *The conductivity matrix M_σ is symmetric positive semi-definite.*

(iii) *The discrete gradient operator G^{red} has full column rank.*

(iv) *The relations $G^{\mathrm{red}} = -\widetilde{S}^{\mathrm{red}\top}$, $\widetilde{G}^{\mathrm{red}} = -S^{\mathrm{red}\top}$ and $C^{\mathrm{red}} = \widetilde{C}^{\mathrm{red}\top}$ hold true.*

(v) *It holds $CG = 0$, $\widetilde{C}\widetilde{G} = 0$, $SC = 0$ and $\widetilde{S}\widetilde{C} = 0$.*

The properties in Assumption 4.20 are considered standard assumptions, that have been also shown and used for instance in [Cle05; Bau12; Cor+19], and are important for the forthcoming numerical analysis. Also note that for 3D discretizations the discrete curl operator C is non-singular, see [Sch11]. For a more detailed insight into the fundamental mathematical questions concerning consistency or convergence of the approximation schemes, see for instance [BK00].

4.6 Conclusions

In this chapter we introduced some fundamentals of classical electromagnetism, namely Maxwell's equations (MEs) and constitutive relations for materials were considered during the work within the *nanoCOPS* project, which are semi-conducting, conducting and isolating materials. In preparation of the spatial discretization and coupling with other devices, we introduced further interface and boundary conditions. Concerning the boundary conditions, the focus was laid upon perfectly electrically conducting boundaries. The very same fundamentals were then transferred to an alternative formulation of MEs, the Lorenz-gauged $\mathbf{A} - \boldsymbol{\varphi}$ formulation (4.21) with a quasi-canonical momentum in order to avoid second order time derivatives of the PDE. This specific approach is similar to the one used by the device modeler `devEM` from the

project partner *MAGWEL* [MAG16] and comes in handy when dealing with semi-conducting material. In order to interface with this software package, the mock element was introduced in the previous chapter.
In Section 4.5 we introduced the finite integration technique (FIT) in order to spatially discretized electromagnetic PDE models, especially the one of consideration (4.21), and incorporated all the boundary conditions, including excitations. Thus we ended up with Maxwell's grid equations in div-grad type Lorenz-gauged $\mathbf{A} - \boldsymbol{\varphi}$ formulation with PEC boundary described by the ODE (4.45), which was already used in [Bau12]. In the accompanying Example 4.10, we provided the FIT for a rectilinear staggered meshes on a brick shaped finite domain, without creating ghost cells. This example is the foundation for a self-implementation of FIT in `python` which was used for the benchmarks in Chapter 7.

5 Coupled Electric Circuits and Electromagnetic Field Devices

In order to satisfy the challenging demands of computer aided design tools in semiconductor industry, especially when integrated circuit dimensions decrease from millimeters to nanometers while their operating frequencies easily exceed the gigahertz range, lumped circuit models alone become insufficient. The mutual electromagnetic influence of neighboring elements, such as cross talking or skin effect, can no longer be covered by equivalent circuits for electromagnetic or semiconductor devices as their number of parameters becomes unreasonably large without physical justification, see for instance [DF08]. This motivates the further need for refined models based on a sufficiently accurate discretization of Maxwell's equations (MEs).

In combination with classical lumped circuit models they provide a great tool for radio-frequency engineers allowing them to test their electromagnetic (EM) field device's design in an realistic work environment instead of describing it by parameter extraction. This was one of the topics we addressed in collaboration with our project partners *ACCO Semiconductor*, *ON Semiconductor* and *NXP Semiconductors* within the EU-founded FP7 project *nanoCOPS* see e. g. [Mat+16; Mat+19]. Similar topics were already addressed by the project's predecessor named *ICESTARS*, see [Bra+11].

This chapter is devoted to coupling approaches for electric circuits, incorporating 3D electromagnetic field models for specific devices, see for example Figure 5.1, that follows the approach in [Bau12], [Sch11] and in the more recent work [ST20]. However, semiconductor devices were also covered in joint project works, see e. g. [Bit+18; Sch+19a].

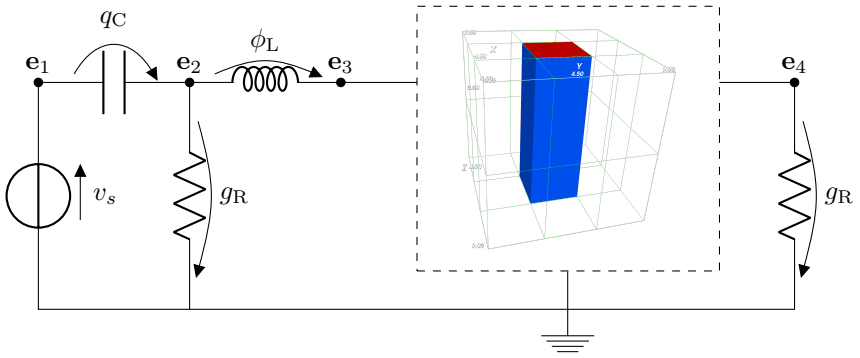


Figure 5.1: Example of an circuit incorporating an EM device.

The so-called field/circuit coupled problems are subjects of many investigations concerning numerical analysis or solving strategies, see e. g. [HM76; Bed93; LM93; De +98; CLP98; DGL99;

Tsu02; DHW04; DW04; Sch+10b; SDW13; CDS19]. Some of them with the focus on quasistatic approximations, such as magneto-quasistatics or electrostatics, where the field equations are spatially discretized using finite-element, finite-difference, boundary-element methods or the finite integration technique (FIT). In this treatise we consider the modeling and analysis of circuits using modified nodal analysis (MNA), see Chapter 3, coupled with full-wave MEs in $\mathbf{A}-\varphi$ formulation which are spatially discretized using FIT, see Chapter 4. The reason for this choice of formulation is that the device electromagnetic modeler `DevEM` of the project partner *MAGWEL* uses this specific one. As `DevEM` also supports semiconductor materials, using the $\mathbf{A}-\varphi$ formulation comes in handy. Further, we use the MNA for the circuit part as most circuit simulation software is pioneered by this modeling approach. In his former works, Baumanns already dealt with this specific kind of coupled systems, see e. g. [Bau12]. We will take this as a basis and provide similar numerical analysis to a variation of different couping approaches whereby we make use of the dissection index concept instead of the tractability index concept.

For the coupling, we persuade a modular approach so that the EM model can be provided by a high grade industry tools such as [MAG16]. It is not unusual, that circuit and EM device simulator are black boxes to each other, see e. g. [Kan01; Zho+06]. Therefore, we describe the EM device as a subcircuit of circuit elements introduced in Section 3 including mock elements. The latter can be done in different ways which we compare against each other when applying waveform relaxation, see Chapter 6. The resulting coupled systems are then analyzed using the dissection index concept and we find a topological index criterion that harmonies with the one given in [Bau12]. To be more precise, the refined model for the EM device under investigation influences the index similar to an inductor. Indeed, it perfectly fits the description of an inductance-like element for the generalized circuit elements concept in [Cor+20]. There it complies with the index statement but for the differentiation index.

The Chapter is organized as follows. First we recall the circuit and EM device subsystems and provide coupling equations which combine the quantities of both. Next, we show how these coupling equations fit into the mock element concept or a combination thereof together with other lumped circuit element introduced in Chapter 3. Followed by this, various equivalent coupled system formulations are gathered and analyzed in terms of a topological index criterion.

5.1 Coupled Modeling

In this section all the model equations are provided in order to incorporate the EM field devices as refined elements into circuits. For simplicity, we assume that there is only one EM device. The here presented approach follows the one in [Bau12; Sch11; ST20].
For the coupled model, the circuit shall be modeled using the MNA, introduced in Chapter 3, and the EM field device using full-wave MEs, see Chapter 4, which are spatially discretized on a finite domain $\Omega \subset \mathbb{R}^3$ using FIT. Further, the EM device's communication with the outside world happens via boundary excitation ϕ_Γ, see (4.41) and current extraction using (4.43). The extraction's and excitation's role can also be interpreted the other way round. We persuade the ansatz to fit the EM device into the mock element's role for later analysis, cf. [ST20]. Note that the terms related to EM devices are decorated with the subscript $(\cdot)_\mathrm{E}$ an those of mock

elements with $(\cdot)_{\mathrm{M}}$. This section, however, only provides the equations and no concrete coupled system formulations which are subject to the next section. For the sake of human readability, the arguments of unknown vector functions will be omitted from here on.

Consider an circuit modeled using the MNA, as introduced in Chapter 3, which consists of standard elements and an additional EM device (E) replacing the mock element (M). Let further Assumption 3.8 be fulfilled. The EM device contributes yet to be defined branch currents $\mathbf{i}_{\mathrm{E}}$ and branch relations A_{E}. Then, the circuit equations read

$$A_{\mathrm{C}} q_{\mathrm{C}}(A_{\mathrm{C}}^{\top}\mathbf{e}, t) + A_{\mathrm{R}} g_{\mathrm{R}}(A_{\mathrm{R}}^{\top}\mathbf{e}, t) + A_{\mathrm{L}}\mathbf{i}_{\mathrm{L}} + A_{\mathrm{V}}\mathbf{i}_{\mathrm{V}} + A_{\mathrm{I}} i_s(t) + A_{\mathrm{E}}\mathbf{i}_{\mathrm{E}} = 0, \tag{5.1a}$$

$$\phi_{\mathrm{L}}(\mathbf{i}_{\mathrm{L}}, t) - A_{\mathrm{L}}^{\top}\mathbf{e} = 0, \tag{5.1b}$$

$$v_s(t) - A_{\mathrm{V}}^{\top}\mathbf{e} = 0 \tag{5.1c}$$

As we can see, we have to define branches for each EM device, or in other words the incidence matrix A_{E} containing the EM devices' branch relations. Then, we are able to define the elements' branch voltages $\mathbf{v}_{\mathrm{E}}$. Further, we have to provide an additional equation for the currents $\mathbf{i}_{\mathrm{E}}$.
In accordance with Section 4.5, let the EM device's model equations be given by Maxwell's grid equations (MGEs) (4.45) meeting Assumption 4.20. Omitting the superscript $(\cdot)^{\mathrm{red}}$ for readability reasons, the model equations read

$$\widetilde{S} M_{\varepsilon} G \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi} + \widetilde{S} M_{\zeta} G M_{\xi} \widetilde{S} M_{\zeta}\mathbf{a} = 0, \tag{5.2a}$$

$$\widetilde{C} M_{\nu} C\mathbf{a} + \frac{\mathrm{d}}{\mathrm{d}t}\big[M_{\varepsilon}(G\boldsymbol{\phi} + G_{\Gamma}\phi_{\Gamma}(t) + \boldsymbol{\pi})\big] + M_{\sigma}\big(G\boldsymbol{\phi} + G_{\Gamma}\phi_{\Gamma}(t) + \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{a}\big) = 0, \tag{5.2b}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{a} - \boldsymbol{\pi} = 0 \tag{5.2c}$$

where ϕ_{Γ} is yet to be defined.

5.1.1 Coupling Branches

As a first step to bring (5.13) and (5.2) together, we have to ensure that the charges are conserved. Recalling that the total flow through the whole boundary adds up to zero, see Lemma 4.1, a compliance with Kirchhoff's circuit laws can only be guaranteed if every part of the boundary is attached to some node of the circuit. Hence, for the coupling we assume that the EM field device is topologically connected to the circuit in terms of the following assumption.

Assumption 5.1
The EM device's boundary $\partial\Omega$ consists of $m_{\mathrm{E}} + 1$ disjoint nonempty parts $\Gamma_j \subset \partial\Omega$, for $j = 0, \dots, m_{\mathrm{E}}$, which are connected to exactly one node of the circuit each. Further, let the ground node be among them which shall be associated with the boundary part Γ_0.

Note that the boundary parts Γ_j must not necessarily be connected sets, for instance they may cover two different contacts regions that are connected to the same circuit node. Further, the

second part of Assumption 5.1 is not really a strong restriction to the design of EM devices since non-contact parts of the boundary or contacts parts that are not supposed to be connected at all, get zero excitation via Dirichlet boundary which represents incidence to the ground node. Nodes, to which the EM device is connected via Γ_j, are the EM device's terminals. By convention, the element's reference terminal is the node attached to Γ_0, that is the ground node. Having it that way is advantageous since the EM device's boundary to node mapping can be expressed in terms of an incidence matrix. The EM device's branches are again defined by the non-reference to reference terminal connections, according to Definition 3.1. Note that the reference terminal is the ground node, for which the according rows in the incidence matrices are eliminated. Therefore, each branch of the EM device has only the non-reference terminal entry in the device's incidence matrix. Since the non-reference terminals are the once connected to Γ_j, for $1 \leq j \leq m_{\mathrm{E}}$, the resulting incidence matrix for the EM device is $A_{\mathrm{E}} \in \{0,1\}^{(n-1)\times m_{\mathrm{E}}}$ with $n-1$ being the total number of circuit nodes but the ground node. In particular

$$(A_{\mathrm{E}})_{i,j} := \begin{cases} 1 & \text{if node } i \text{ is a non-reference terminal of branch } j \text{ / is connected to } \Gamma_j, \\ 0 & \text{else.} \end{cases}$$

Note that with the convention that Γ_0 is connected to the ground node, this incidence matrix may be different to the one considered in [Bau12]. With the incidence matrix given, the branch voltages $\mathbf{v}_{\mathrm{E}} : \mathcal{I} \to \mathbb{R}^{m_{\mathrm{E}}}$ are obtained by the expression $\mathbf{v}_{\mathrm{E}} = A_{\mathrm{E}}^{\top}\mathbf{e}$.

As a next step of preparation, we recall the geometric objects of the EM device resulting from its spatial discretization in Section 4.5 using two staggered grids. These objects were collected respectively in the primal and dual cell complexes $\mathcal{G} = \mathcal{P} \cup \mathcal{L} \cup \mathcal{F} \cup \mathcal{V}$ and $\widetilde{\mathcal{G}} = \widetilde{\mathcal{P}} \cup \mathcal{L} \cup \mathcal{F} \cup \mathcal{V}$ which consist of mesh points, links and volumes each. The primal mesh points belonging to the boundary are $P(i)$, for $i \in \mathcal{N}_{\mathcal{P}}^{\mathrm{bound}}$, and the primal intern links are $L(i)$, for $i \in \mathcal{N}_{\mathcal{L}}^{\mathrm{int}}$. The dual volumes incident to the boundary $\partial\Omega$ are $\widetilde{V}(i)$, for $i \in \mathcal{N}_{\mathcal{P}}^{\mathrm{bound}}$.

In order to avoid contact regions containing solely isolated mesh points and being inconsistent with the spatial discretization, we require the following assumption.

Assumption 5.2

(i) Let each contact region Γ_j contain at least one primal boundary mesh point $P(i)$ with $i \in \mathcal{N}_{\mathcal{P}}^{\mathrm{bound}}$ that is incident to a primal link $L(k)$ with $k \in \mathcal{N}_{\mathcal{L}}^{\mathrm{int}}$, see Section 4.5.6.

(ii) For all $P(i)$ belonging to the contact region Γ_j holds $i \in \mathcal{N}_{\mathcal{P}}^{\mathrm{bound}}$ and $\partial\widetilde{V}(i) \cap \Omega \subset \bar{\Gamma}_j$.

5.1.2 Coupling Potentials

In order to match circuit variables with the boundary excitation of the EM device, consider the following matrix

$$\Lambda \in \{0,1\}^{\mathcal{N}_{\mathcal{P}}^{\mathrm{bound}} \times m_{\mathrm{E}}}, \qquad (\Lambda)_{i,j} = \begin{cases} 1 & \text{if mesh point } P(i) \text{ belongs to } \Gamma_j \\ 0 & \text{else} \end{cases}.$$

Then, with ΛA_{E} we obtain a non-reference terminal to primal boundary mesh point mapping. Note that the EM device's reference terminal is not mapped since it is the circuit's ground node, as of Assumption 5.1. In other words, for the reference terminal there is no variable to map and the excitation at Γ_0 is set to zero by the definition of Λ. As a consequence, the incident node potentials are the branch voltages. Therefore, we match the EM device's terminal potentials with the corresponding discrete electric scalar potential excitation ϕ_Γ of the EM field device by the equation

$$\phi_\Gamma = \Lambda A_{\mathrm{E}}^\top \mathbf{e}. \tag{5.3}$$

Equation (5.3) mimics the continuous physical contact case (4.14). In terms of the EM device's branch voltages, the excitation coupling reads

$$\phi_\Gamma = \Lambda \mathbf{v}_{\mathrm{E}}.$$

We obtain the following mapping property concerning Λ.

Lemma 5.3
Given Assumptions 5.1 and 5.2, the matrix $G_\Gamma \Lambda$ has full column rank.

Proof. As each region Γ_j, for $j = 1, \ldots, m_{\mathrm{E}}$, contains at least one boundary mesh point that is not shared by the other regions, Λ has full column rank and maximal one nonzero entry per row. Note that also G_Γ, by construction in Section 4.5, has exactly one nonzero entry in each row belonging to a primal link that has an incident primal boundary mesh point, otherwise the row is zero. To be more precise, if the link is incident, it has either -1 or 1 in the column belonging to that boundary mesh point. For $G_\Gamma \Lambda$ it is now guaranteed, by Assumption 5.2, that each column of Λ has at least on row entry, that hits such an entry in G_Γ resulting in a nontrivial column. Since for any other column in Λ, the same nonzero entries in G_Γ are never hit again, as the regions do not share the same primal boundary mesh points as of Assumption 5.1, the resulting columns have nonzero entries in different rows only. Therefore, for each column of $G_\Gamma \Lambda$, there exists a row that has only for that column a nonzero entry and consequently $G_\Gamma \Lambda$ has full column rank. □

5.1.3 Coupling Currents

For the branch currents $\mathbf{i}_{\mathrm{E}} : \mathcal{I} \to \mathbb{R}^{m_{\mathrm{E}}}$, we derive three different expressions which are (5.10), (5.11) and (5.12). The first variant is the one already considered by Baumanns in [Bau12] on which the here presented analysis is partially building upon. The second and third variants are motivated by industrial usage of which, for instance, the device electromagnetic modeler `DevEM` makes use of to interface with electric circuit modeler [Sch+16].

We start by recalling the model for total current density $\mathbf{J}_{\mathrm{tot}} = \mathbf{J} + \frac{\partial}{\partial t}\mathbf{D}$ as introduced in Chapter 4, see equation (4.4). With this, the total current through a surface Γ is given by

$$i_\Gamma = \int_\Gamma \mathbf{J}_{\mathrm{tot}} \cdot \mathrm{d}s = \int_\Gamma \left[\frac{\partial}{\partial t}\mathbf{D} + \mathbf{J}\right] \cdot \mathrm{d}s,$$

see for instance (4.15). After spatial discretization and in terms potentials with incorporated boundary excitation, the total current flow is to be understood as a quantity living on a dual facets $\widetilde{F}(k)$, for $k \in \mathcal{N}_{\mathcal{L}}$, given by (4.43), that is

$$\left(\widetilde{C} M_\nu C \bar{\mathbf{a}}\right)_k$$

where the superscript $(\cdot)^{\mathrm{red}}$ was removed for consistent notation. According to Assumption 5.2, the surface of a contact region Γ_j, for $j = 0, \dots, m_{\mathrm{E}}$, is given by the union

$$\bigcup_{P(i) \subset \Gamma_j} \partial \widetilde{V}(i) \cap \partial\Omega.$$

Making use of the discrete operators properties (4.29) and (4.32) we observe, that the total current flow $i_{\Gamma,i}$ through $\partial \widetilde{V}(i) \cap \partial\Omega$, for $i \in \mathcal{N}_{\mathcal{P}}$ can be expressed in terms of the volume's divergence through the complementary part, that is

$$i_{\Gamma,i} = \sum_{\widetilde{F}(k) \subset \widetilde{V}(i) \setminus \partial\Omega} \widetilde{\zeta}^{[3]}(\widetilde{F}(k), \widetilde{V}(i)) (\widetilde{C} M_\nu C \bar{\mathbf{a}})_k \tag{5.4}$$

where the orientation is corrected by the window function $\widetilde{\zeta}^{[3]}$ according to Definition 4.14. By the discrete operators' duality and the incorporation of boundary conditions, the expression (5.4) can be summarized by

$$i_\Gamma = \widetilde{S}_\Gamma \widetilde{C} M_\nu C \bar{\mathbf{a}} = -G_\Gamma^\top \widetilde{C} M_\nu C \bar{\mathbf{a}}$$

for $i_\Gamma = (i_{\Gamma,i})_{i \in \mathcal{N}_{\mathcal{P}}^{\mathrm{bound}}}$. The branch currents are then defined by the sum of all total currents through $i_{\Gamma,i}$ with $P(i) \subset \Gamma_j$ that is

$$\mathbf{i}_{\mathrm{E}} = -\Lambda^\top G_\Gamma^\top \widetilde{C} M_\nu C \mathbf{a} \tag{5.5}$$

which was already introduced in [Bau12]. As the div-curl identity is a property the discrete operators inherit from the continuous case, see the properties in Assumption 4.20, and due to Assumption 5.1, the reference terminal currents are obtained as the negative sum of the non-reference terminal currents and hence complying with KCL.

According to the discrete MA in (5.2b), we can also switch

$$\widetilde{C} M_\nu C \mathbf{a} \quad \text{with} \quad \frac{\mathrm{d}}{\mathrm{d}t} \left(M_\varepsilon \left(G\boldsymbol{\phi} + G_\Gamma \phi_\Gamma(t) + \boldsymbol{\pi} \right) \right) + M_\sigma \left(G\boldsymbol{\phi} + G_\Gamma \phi_\Gamma(t) + \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{a} \right)$$

yielding the alternative representation of the branch currents

$$\mathbf{i}_{\mathrm{E}} = \Lambda^\top G_\Gamma^\top \Big[\frac{\mathrm{d}}{\mathrm{d}t} \big(M_\varepsilon (G\boldsymbol{\phi} + G_\Gamma \phi_\Gamma(t) + \boldsymbol{\pi}) \big) + M_\sigma \big(G\boldsymbol{\phi} + G_\Gamma \phi_\Gamma(t) + \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{a} \big) \Big]. \tag{5.6}$$

Further exploitation of the canonical momentum, that is replacing $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{a}$ with $\boldsymbol{\pi}$ as legitimated by (5.2c), yields the third analytically equivalent variant

$$\mathbf{i}_{\mathrm{E}} = \Lambda^\top G_\Gamma^\top \Big[\frac{\mathrm{d}}{\mathrm{d}t} \big(M_\varepsilon (G\boldsymbol{\phi} + G_\Gamma \phi_\Gamma(t) + \boldsymbol{\pi}) \big) + M_\sigma \big(G\boldsymbol{\phi} + G_\Gamma \phi_\Gamma(t) + \boldsymbol{\pi} \big) \Big]. \tag{5.7}$$

As already stated, the latter equivalent alternatives are used by industrial tools which motivates the comparison of arising coupled system formulations by means of numerical methods which we cover in the latter chapters. To be more precise, there are formulations of coupled systems which fail to be solved using waveform relaxations (Section 6.3) whereby others do not and vice versa, see for instance Example 7.7 in Chapter 7. The here considered formulations are given in the following section.

5.2 Coupled System Formulations

In the following, we derive three equivalent classes of field/circuit coupled system formulations which are denoted by incorporated, shifted and black-box versions. Each variant provides a different form resulting from various approaches of how the EM device's model (5.2) fits the mock element description (3.5) of Chapter 3 by means of constitutive equations derived by the previously introduced coupling equations (5.3), (5.5), (5.6) and (5.7). In [Sch11; Alì+12], for instance, similar coupling approaches were investigated but for a different physical problem. There, terms such as source or parameter coupling are introduced, to which parallels can be drawn.
The first class of formulations, that is the incorporated versions (5.21)-(5.23), results from interpreting the EM device as a mock element using respectively the constitutive equations (5.10), (5.11) and (5.12). An example is shown in Figure 5.2.
The class of shifted versions contains the two formulations (5.24) and (5.25). They arise from interpreting circuit variable related terms of the current coupling equations in (5.16) and (5.17) as generalized capacitors, see further [Cor+20], whereby the remaining terms define the constitutive equations to interpret the EM device as an other mock element, see Figure 5.3. Note that such an interpretation cannot be deduced by the current coupling equation (5.5) due to the lack of circuit variables. The latter class of formulations contains the black-box versions which are (5.26)-(5.28). They differ from the incorporated ones only in that the constitutive equations (5.10), (5.11) and (5.12) are not incorporated into the circuit equations.

Although all formulations are analytically equivalent, see Proposition 5.10, they divide into two equivalence classes when applied either Gauss-Seidel or Jacobi type waveform relaxation schemes to them, see Corollaries 6.15 and 6.17 in Section 6.3. Moreover, formulations across these equivalence classes lead to different convergence behavior which is exemplary shown in Chapter 7.

5.2.1 EM Device as Mock Element

For the EM device to fit the mock element introduced in Chapter 3, we have to provide a constitutive equation (3.5) and an intrinsic equation (3.6). For this we make use of the previously introduced coupled modeling equations, leading to different mock element interpretations, collected in Lemmata 5.6 and 5.8.

Let again be $\mathbf{x} = (\mathbf{e}, \mathbf{i}_\mathrm{L}, \mathbf{i}_\mathrm{V})$ and $\mathbf{u} = (\boldsymbol{\phi}, \mathbf{a}, \boldsymbol{\pi})$. Further, with $\mathbf{w} = (\mathbf{w}_\mathrm{C}, \mathbf{w}_\mathrm{L}, \mathbf{w}_\mathrm{E})$ and $\mathbf{w}_\mathrm{u} = (\mathbf{w}_\phi, \mathbf{w}_\mathrm{a}, \mathbf{w}_\pi)$ we denote the arguments that are placeholder for derivative terms. Again we omit the unknowns argument for readability reasons.

We start by incorporating the excitation resulting from the circuit (5.3) into the EM device's model equations (5.2) yielding the system

$$M_\mathrm{E}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_\mathrm{E}(\mathbf{u}) = c_\mathrm{E}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}), \mathbf{x}) \tag{5.8}$$

where

$$M_\mathrm{E} := \begin{bmatrix} \widetilde{S}M_\varepsilon G & 0 & 0 \\ M_\varepsilon G & M_\sigma & M_\varepsilon \\ 0 & I & 0 \end{bmatrix}, \qquad b_\mathrm{E}(\mathbf{u}) := \begin{bmatrix} 0 & \widetilde{S}M_\zeta G M_\xi \widetilde{S} M_\zeta & 0 \\ M_\sigma G & \widetilde{C}M_\nu C & 0 \\ 0 & 0 & -I \end{bmatrix} \mathbf{u}$$

and the circuit variables are collected on the right-hand side by

$$c_\mathrm{E}(\mathbf{w}_\mathrm{E}, \mathbf{x}) := -\begin{pmatrix} 0 \\ \mathbf{w}_\mathrm{E} + M_\sigma G_\Gamma \Lambda A_\mathrm{E}^\top \mathbf{e} \\ 0 \end{pmatrix}, \qquad d_{\mathrm{E-MNA}}(\mathbf{x}) := M_\varepsilon G_\Gamma \Lambda A_\mathrm{E}^\top \mathbf{e}. \tag{5.9}$$

Lemma 5.4
Provided Assumption 4.20 is fulfilled, the matrix M_E in (5.8) is nonsingular with the inverse

$$M_\mathrm{E}^{-1} = \begin{bmatrix} (\widetilde{S}M_\varepsilon G)^{-1} & 0 & 0 \\ 0 & 0 & I \\ -G(\widetilde{S}M_\varepsilon G)^{-1} & M_\varepsilon^{-1} & -M_\varepsilon^{-1}M_\sigma \end{bmatrix}.$$

Proof. As of Assumption 4.20, the discrete gradient operator G has full column rank. Since further $\widetilde{S} = -G^\top$ and M_ε is positive definite, $G^\top M_\varepsilon G$ is positive definite with an existing inverse. Thus, the expressions $(\widetilde{S}M_\varepsilon G)^{-1} = -(G^\top M_\varepsilon G)^{-1}$ and M_ε^{-1} are well-defined and consequently M_E^{-1}. From $M_\mathrm{E}M_\mathrm{E}^{-1} = I$ follows the statement. □

Lemma 5.5
The EM device's functions b_E and c_E are Lipschitz continuous as of Definition A.10.

Proof. This is an immediate consequence of the functions linearity by matrix-matrix and matrix-vector multiplications. □

Next, we consider the current coupling equations. Whereas (5.5) is only dependent on the EM field variables, the other two current coupling equations (5.6) and (5.7) are additionally dependent on circuit variables which becomes clear when inserting the excitation equation (5.3) into them. The constitutive relation for the mock element obtained from (5.5) reads

$$\mathbf{i}_\mathrm{E} = f_{\mathrm{E}0}(\mathbf{u}) := -\Lambda^\top G_\Gamma^\top \widetilde{C} M_\nu C \mathbf{a}. \tag{5.10}$$

Contrary to this variant, we obtain from inserting (5.3) in (5.6)

$$\mathbf{i}_{\mathrm{E}} = f_{\mathrm{E1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{x}, \mathbf{u}) \tag{5.11}$$

with

$$f_{\mathrm{E1}}(\mathbf{w}_{\mathrm{E}}, \mathbf{w}_{\mathrm{u}}, \mathbf{x}, \mathbf{u}) := \Lambda^{\top} G_{\Gamma}^{\top} \left[M_{\varepsilon} \left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi} \right) + \mathbf{w}_{\mathrm{E}} + M_{\sigma} \left(G\boldsymbol{\phi} + G_{\Gamma} \Lambda A_{\mathrm{E}}^{\top} \mathbf{e} + \mathbf{w}_{\mathrm{a}} \right) \right].$$

Note that shifting the differentiation operator in front of each component of $\mathbf{u}$ does not require any further smoothness since M_{ε} is nonsingular and G has full column rank as of Assumption 4.20. Similarly, inserting (5.3) in (5.7) yields the last variant of constitutive equation

$$\mathbf{i}_{\mathrm{E}} = f_{\mathrm{E2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{x}, \mathbf{u}) \tag{5.12}$$

with

$$f_{\mathrm{E2}}(\mathbf{w}_{\mathrm{E}}, \mathbf{w}_{\mathrm{u}}, \mathbf{x}, \mathbf{u}) := \Lambda^{\top} G_{\Gamma}^{\top} \left[M_{\varepsilon} \left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi} \right) + \mathbf{w}_{\mathrm{E}} + M_{\sigma} \left(G\boldsymbol{\phi} + G_{\Gamma} \Lambda A_{\mathrm{E}}^{\top} \mathbf{e} + \boldsymbol{\pi} \right) \right].$$

Hence, we obtaine three possibilities to express the EM device as a mock element in terms of (3.5). We summarize these observation by the following Lemma.

Lemma 5.6 (EM device as mock element)
The circuit (5.1) *incorporating the EM device* (5.2) *can be expressed by the circuit*

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}}(\mathbf{x}, t), \mathbf{x}, t) = c_{\mathrm{MNA}}(\mathbf{i}_{\mathrm{M}}) \tag{5.13}$$

where

$$f_{\mathrm{MNA1}}(\mathbf{w}, \mathbf{x}, t) = \begin{pmatrix} A_{\mathrm{C}}\mathbf{w}_{\mathrm{C}} + A_{\mathrm{R}} g_{\mathrm{R}}(A_{\mathrm{R}}^{\top}\mathbf{e}, t) + A_{\mathrm{L}}\mathbf{i}_{\mathrm{L}} + A_{\mathrm{V}}\mathbf{i}_{\mathrm{V}} + A_{\mathrm{I}} i_s(t) \\ \mathbf{w}_{\mathrm{L}} - A_{\mathrm{L}}^{\top}\mathbf{e} \\ v_s(t) - A_{\mathrm{V}}^{\top}\mathbf{e} \end{pmatrix},$$

$$d_{\mathrm{MNA1}}(\mathbf{x}, t) = \begin{pmatrix} q_{\mathrm{C}}(A_{\mathrm{C}}^{\top}\mathbf{e}, t) \\ \phi_{\mathrm{L}}(\mathbf{i}_{\mathrm{L}}, t) \end{pmatrix}, \qquad c_{\mathrm{MNA}}(\mathbf{i}_{\mathrm{M}}) = \begin{pmatrix} -A_{\mathrm{M}}\mathbf{i}_{\mathrm{M}} \\ 0 \\ 0 \end{pmatrix}.$$

and the mock element is given by

(i) $A_{\mathrm{M}} = A_{\mathrm{E}}$, $\mathbf{i}_{\mathrm{M}} = f_{\mathrm{E0}}(\mathbf{u})$ *and* (5.8).

(ii) $A_{\mathrm{M}} = A_{\mathrm{E}}$, $\mathbf{i}_{\mathrm{M}} = f_{\mathrm{E1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{x}, \mathbf{u})$ *and* (5.8).

(iii) $A_{\mathrm{M}} = A_{\mathrm{E}}$, $\mathbf{i}_{\mathrm{M}} = f_{\mathrm{E2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{x}, \mathbf{u})$ *and* (5.8).

Proof. By Lemma 5.4 M_{E} is nonsingular so that we can define $M_{\mathrm{M}} := M_{\mathrm{E}}$. Since the functions c_{E} and $d_{\mathrm{E-MNA}}$ only make use of the expression $A_{\mathrm{E}}^{\top}\mathbf{e}$, as of definition in (5.9), we can find function d_{M} and c_{M} as follows

$$d_{\mathrm{M}}(A_{\mathrm{M}}^{\top}\mathbf{e}) := d_{\mathrm{E-MNA}}(\mathbf{x}),$$
$$c_{\mathrm{M}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{M}}(A_{\mathrm{M}}^{\top}\mathbf{e}), A_{\mathrm{M}}^{\top}\mathbf{e}) := c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{M}}(A_{\mathrm{M}}^{\top}\mathbf{e}), \mathbf{x})$$

fitting the mock element's intrinsic description (3.6). Concerning (i), we can simply set

$$f_{\mathrm{M}}(d_{\mathrm{M}}(A_{\mathrm{M}}^{\top}\mathbf{e}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}_{\mathrm{M}}, A_{\mathrm{M}}^{\top}\mathbf{e}, \mathbf{u}_{\mathrm{M}}) := f_{\mathrm{E0}}(\mathbf{u}_{\mathrm{M}})$$

whereby $\mathbf{u}_{\mathrm{M}} = \mathbf{u}$ fitting the mock element's constitutive description (3.5). For (ii) we notice that from the third argument of f_{E1} only $\mathbf{e}$ is used and immediately multiplied by $A_{\mathrm{E}}^{\top}$ as of definition (5.11). Therefore, we are allowed to define

$$f_{\mathrm{M}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{M}}(A_{\mathrm{M}}^{\top}\mathbf{e}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}_{\mathrm{M}}, A_{\mathrm{M}}^{\top}\mathbf{e}, \mathbf{u}_{\mathrm{M}}) := f_{\mathrm{E1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{M}}(A_{\mathrm{M}}^{\top}\mathbf{e}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}_{\mathrm{M}}, \mathbf{x}, \mathbf{u}_{\mathrm{M}}),$$

fitting the mock element's description (3.5). Analogously, (iii) fits the mock element's description (3.5). □

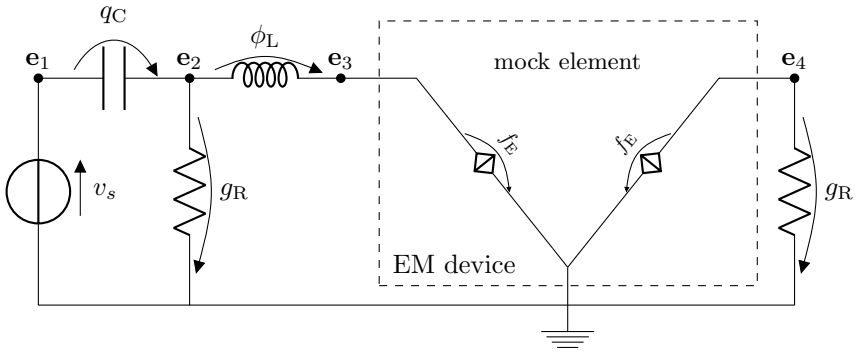


Figure 5.2: Circuit incorporating an EM device entirely as a mock element, here represented by controlled sources along each branch with $f_{\mathrm{E}} \in \{f_{\mathrm{E0}}, f_{\mathrm{E1}}, f_{\mathrm{E2}}\}$.

As mentioned above, the variants (5.11) and (5.12) enable the possibility to shift terms concerning circuit variables into the circuit equation since these terms comply with the form of capacitance-like elements. Consequently, the circuit would be a different one but so the coupling equation. In order to do so, the summands in f_{E1} and f_{E2} are split up by introducing new functions

$$\begin{aligned}
q_{\mathrm{E}}(A_{\mathrm{E}}^{\top}\mathbf{e}) &:= \Lambda^{\top} G_{\Gamma}^{\top} M_{\varepsilon} G_{\Gamma} \Lambda A_{\mathrm{E}}^{\top}\mathbf{e}, \\
g_{\mathrm{E}}(A_{\mathrm{E}}^{\top}\mathbf{e}) &:= \Lambda^{\top} G_{\Gamma}^{\top} M_{\sigma} G_{\Gamma} \Lambda A_{\mathrm{E}}^{\top}\mathbf{e}, \\
s_{\mathrm{E1}}(\mathbf{w}_{\mathrm{u}}, \mathbf{u}) &:= \Lambda^{\top} G_{\Gamma}^{\top} \left[M_{\varepsilon}\left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi}\right) + M_{\sigma}\left(G\boldsymbol{\phi} + \mathbf{w}_{\mathrm{a}}\right)\right] && (5.14) \\
s_{\mathrm{E2}}(\mathbf{w}_{\mathrm{u}}, \mathbf{u}) &:= \Lambda^{\top} G_{\Gamma}^{\top} \left[M_{\varepsilon}\left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi}\right) + M_{\sigma}\left(G\boldsymbol{\phi} + \boldsymbol{\pi}\right)\right] && (5.15)
\end{aligned}$$

yielding that (5.11) is equivalent to

$$\mathbf{i}_{\mathrm{E}} = \frac{\mathrm{d}}{\mathrm{d}t} q_{\mathrm{E}}(A_{\mathrm{E}}^{\top}\mathbf{e}) + g_{\mathrm{E}}(A_{\mathrm{E}}^{\top}\mathbf{e}) + s_{\mathrm{E1}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u}) \tag{5.16}$$

and (5.12) is equivalent to

$$\mathbf{i}_{\mathrm{E}} = \frac{\mathrm{d}}{\mathrm{d}t} q_{\mathrm{E}}(A_{\mathrm{E}}^{\top}\mathbf{e}) + g_{\mathrm{E}}(A_{\mathrm{E}}^{\top}\mathbf{e}) + s_{\mathrm{E2}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u}). \tag{5.17}$$

Lemma 5.7
Given Assumptions 4.20, 5.1 and 5.2, the function q_E is strongly monotone.

Proof. By Assumption 4.20, M_ε is positive definite. With Assumptions 5.1 and 5.2 to hold, we can make use of Lemma 5.3 stating that $G_\Gamma \Lambda$ has full column rank. Consequently, $\Lambda^\top G_\Gamma^\top M_\varepsilon G_\Gamma \Lambda$ is positive definite and q_E is strongly monotone as of Lemma A.14. □

For both constitutive element equations (5.16) and (5.17), shifting the q_E and q_E terms into the MNA allows us to interpret the EM device as a composition of capacitors and possible resistors, alongside each branch of another mock element, defined by the remaining term, see Figure 5.3. In fact, the parallel capacitors with possible resistors form, in itself, a capacitance-like element, see [Cor+20]. A similar composition with a capacitive part extraction has been presented in [Alì+12] for the coupling of circuit and spatially discretized semiconductor device models or with inductive part as optimized transmission condition in [Cor+17; Cor20] enabling a connection to Schwarz methods.
Following this idea, we group the constitutive equations of capacitors and resistors with the according parts of either expression by introducing

$$A_\mathcal{C} := \begin{bmatrix} A_\mathrm{C} & A_\mathrm{E} \end{bmatrix}, \qquad A_\mathcal{R} := \begin{bmatrix} A_\mathrm{R} & A_\mathrm{E} \end{bmatrix}, \qquad q_\mathcal{C} := \begin{pmatrix} q_\mathrm{C} \\ q_\mathrm{E} \end{pmatrix}, \qquad g_\mathcal{R} := \begin{pmatrix} g_\mathrm{R} \\ g_\mathrm{E} \end{pmatrix}.$$

Lemma 5.8 (EM device as different mock element)
The circuit (5.1) incorporating the EM device (5.2) can be expressed by the topologically modified circuit

$$f_\mathrm{MNA2}(\frac{\mathrm{d}}{\mathrm{d}t} d_\mathrm{MNA2}(\mathbf{x}, t), \mathbf{x}, t) = c_\mathrm{MNA}(\mathbf{i}_\mathrm{M}) \tag{5.18}$$

where

$$f_\mathrm{MNA2}(\mathbf{w}, \mathbf{x}, t) := \begin{pmatrix} A_\mathcal{C}\mathbf{w}_\mathrm{C} + A_\mathcal{R} g_\mathcal{R}(A_\mathcal{R}^\top \mathbf{e}, t) + A_\mathrm{L}\mathbf{i}_\mathrm{L} + A_\mathrm{V}\mathbf{i}_\mathrm{V} + A_\mathrm{I} i_s(t) \\ \mathbf{w}_\mathrm{L} - A_\mathrm{L}^\top \mathbf{e} \\ v_s(t) - A_\mathrm{V}^\top \mathbf{e} \end{pmatrix},$$

$$d_\mathrm{MNA2}(\mathbf{x}) := \begin{pmatrix} q_\mathcal{C}(A_\mathcal{C}^\top \mathbf{e}, t) \\ \phi_\mathrm{L}(\mathbf{i}_\mathrm{L}, t) \end{pmatrix}, \qquad c_\mathrm{MNA}(\mathbf{i}_\mathrm{M}) = \begin{pmatrix} -A_\mathrm{M}\mathbf{i}_\mathrm{M} \\ 0 \\ 0 \end{pmatrix}.$$

and the mock element is given by

(i) $A_\mathrm{M} = A_\mathrm{E}$, $\mathbf{i}_\mathrm{M} = s_\mathrm{E1}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u})$ *and* (5.8).

(ii) $A_\mathrm{M} = A_\mathrm{E}$, $\mathbf{i}_\mathrm{M} = s_\mathrm{E2}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u})$ *and* (5.8).

Proof. In the proof ol Lemma 5.6 we have already shown that (5.8) fits the mock element's intrinsic description (3.6). Further, for (i) we notice that s_E1 is independent of circuit variables hence we can define

$$f_\mathrm{M}(d_\mathrm{M}(A_\mathrm{M}^\top \mathbf{e}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}_\mathrm{M}, A_\mathrm{M}^\top \mathbf{e}, \mathbf{u}_\mathrm{M}) := s_\mathrm{E1}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}_\mathrm{M}, \mathbf{u}_\mathrm{M})$$

fitting the mock element's description (3.5). Analogously, (ii) fits the mock element's description (3.5). □

An example of how the EM device is to be understood as a mock element in accordance with Lemma 5.8 is given in Figure 5.3.

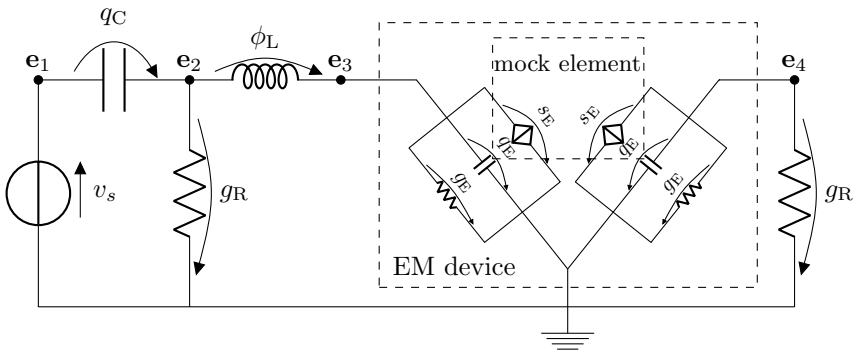


Figure 5.3: Circuit incorporating an EM device whose constitutive element equations are interpreted as a composition of capacitors, resistors (possibly degenerated and hence positive semi-definite) and a mock element, here represented by controlled sources along each branch with $s_E \in \{s_{E1}, s_{E2}\}$.

Interpreting the EM device as a mock element in terms of Lemma 5.6 yields, respectively, for (5.13), the new right hand sides

$$c_{\mathrm{MNA1.0}}(\mathbf{u}), \tag{5.19a}$$

$$c_{\mathrm{MNA1.1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{x}, \mathbf{u}) \quad \text{or} \tag{5.19b}$$

$$c_{\mathrm{MNA1.2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{x}, \mathbf{u}) \tag{5.19c}$$

with

$$c_{\mathrm{MNA1.0}}(\mathbf{u}) := \begin{pmatrix} A_{\mathrm{E}}\Lambda^\top G_\Gamma^\top \widetilde{C} M_\nu C\mathbf{a} \\ 0 \\ 0 \end{pmatrix},$$

$$c_{\mathrm{MNA1.1}}(\mathbf{w}_{\mathrm{E}}, \mathbf{w}_{\mathrm{u}}, \mathbf{x}, \mathbf{u}) := \begin{pmatrix} -A_{\mathrm{E}}\Lambda^\top G_\Gamma^\top \left[M_\varepsilon \left(G\mathbf{w}_\phi + \mathbf{w}_\pi\right) + \mathbf{w}_{\mathrm{E}} + M_\sigma \left(G\boldsymbol{\phi} + G_\Gamma \Lambda A_{\mathrm{E}}^\top \mathbf{e} + \mathbf{w}_{\mathrm{a}}\right)\right] \\ 0 \\ 0 \end{pmatrix},$$

$$c_{\mathrm{MNA1.2}}(\mathbf{w}_{\mathrm{E}}, \mathbf{w}_{\mathrm{u}}, \mathbf{x}, \mathbf{u}) := \begin{pmatrix} -A_{\mathrm{E}}\Lambda^\top G_\Gamma^\top \left[M_\varepsilon \left(G\mathbf{w}_\phi + \mathbf{w}_\pi\right) + \mathbf{w}_{\mathrm{E}} + M_\sigma \left(G\boldsymbol{\phi} + G_\Gamma \Lambda A_{\mathrm{E}}^\top \mathbf{e} + \boldsymbol{\pi}\right)\right] \\ 0 \\ 0 \end{pmatrix}.$$

In order to realize the shifting approach of Lemma 5.8, we introduce the new right-hand sides for (5.18)

$$c_{\mathrm{MNA2.1}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u}) \quad \text{or} \tag{5.20a}$$

$$c_{\mathrm{MNA2.2}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u}) \tag{5.20b}$$

with

$$c_{\mathrm{MNA2.1}}(\mathbf{w}_{\mathrm{u}}, \mathbf{u}) := \begin{pmatrix} -A_{\mathrm{E}}\Lambda^{\top} G_{\Gamma}^{\top} \left[M_{\varepsilon} \left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi} \right) + M_{\sigma} \left(G\boldsymbol{\phi} + \mathbf{w}_{\mathrm{a}} \right) \right] \\ 0 \\ 0 \end{pmatrix},$$

$$c_{\mathrm{MNA2.2}}(\mathbf{w}_{\mathrm{u}}, \mathbf{u}) := \begin{pmatrix} -A_{\mathrm{E}}\Lambda^{\top} G_{\Gamma}^{\top} \left[M_{\varepsilon} \left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi} \right) + M_{\sigma} \left(G\boldsymbol{\phi} + \boldsymbol{\pi} \right) \right] \\ 0 \\ 0 \end{pmatrix}.$$

5.2.2 Resulting Coupled Systems

A coupled electric circuit and electromagnetic field device model is obtained by joining the circuit model (3.8) of Section 3 and the MGEs (4.45) from Section 4, for each EM field device, using the just provided subsystem representations and constitutive equations. The resulting coupled model's system of differential-algebraic equations (DAE) describes an ODE-DAE additive coupled system. In rare cases, that is if the there is a C-tree spanning the circuit, see Theorem 5.12, the coupled model is a system of ODEs describing an ODE-ODE additive coupled system.

The expressions listed above already lead to a variety of combinations. In particular, taking the EM subsystem (5.8), which is a fixed subsystem among all variations, we can incorporated it as a mock element into the circuit, such as (5.13) or as a different mock element in a modified circuit, that is (5.18). As for the first variant, we use concatenations of either (5.10), (5.11) or (5.12), for the mock elements' characteristic functions, that is realized by right hand sides of from (5.19). Similarly, taking the modified circuit, the constitutive equations' residuals for $\mathbf{i}_{\mathrm{E}}$, see Lemma 5.8, are then used for the mock elements' characteristic functions yielding right-hand sides of the from (5.20).
On the basis of the above preliminaries, we now provide three categories of system formulations for field/circuit coupled problems. These categories aim to help with design decisions for an interface that allows commercial circuit simulators to communicate with EM device modeler, such as `DevEM` as part of the *nanoCOPS* project. The versions of the first category aim for a natural incorporation of the refined models' constitutive equations into the MNA as current controlling elements. These versions are used during recent simulations of field/circuit coupled problems using either a self-implemented version of FIT or the model equations from `DevEM` for the EM devices' subsystems. The second category provides forms where the circuit and EM device terms are clearly separated. These versions are so far in an experimental stage and are only possible with the self-implemented version of FIT and linear `DevEM` models that are extracted from the current available interface. The latter category contains versions which were used in an early stage of the interface design for the communication with `DevEM`, e. g. in [Sch+16]. Whereas all the system formulations are analytically equivalent, these categories are either motivated by implementation purposes or by their interesting structure for the forthcoming solving processes, see Section 6.

Let $\mathcal{I} = [t_0, T] \subset \mathbb{R}$ be a time interval. Further, all forthcoming additive coupled systems shall comply the initial conditions $\mathbf{x}(t_0) = x_0$ and $\mathbf{u}(t_0) = u_0$, and if needed $\mathbf{i}_{\mathrm{E}}(t_0) = i_{\mathrm{E}0}$, for consistent initial values x_0, u_0 and $i_{\mathrm{E}0}$.

Incorporated Versions Following the idea of the MNA approach to eliminate as many current quantities as possible, we incorporate the current coupling equation into KCL of the circuit's subsystem, i. e. using any right hand side of (5.19) for (5.13). The resulting field/circuit coupled systems can be represented by the forms

$$M_{\mathrm{E}} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_{\mathrm{E}}(\mathbf{u}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \mathbf{x}), \tag{5.21a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}}(\mathbf{x}), \mathbf{x}, t) = c_{\mathrm{MNA1.0}}(\mathbf{u}), \tag{5.21b}$$

$$M_{\mathrm{E}} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_{\mathrm{E}}(\mathbf{u}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \mathbf{x}), \tag{5.22a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}}(\mathbf{x}), \mathbf{x}, t) = c_{\mathrm{MNA1.1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{x}, \mathbf{u}), \tag{5.22b}$$

$$M_{\mathrm{E}} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_{\mathrm{E}}(\mathbf{u}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \mathbf{x}), \tag{5.23a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}}(\mathbf{x}), \mathbf{x}, t) = c_{\mathrm{MNA1.2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{x}, \mathbf{u}) \tag{5.23b}$$

which are called *incorporated versions*. Even though not directly visible, the latter two are indeed as well additive coupled systems by Definition 2.9 since $c_{\mathrm{MNA1.1}}$ and $c_{\mathrm{MNA1.2}}$ are linear in all components as of (5.19), but not yet brought in the form of (2.7). The reason for this lies within the communication regime of the subsystems which we address in the forthcoming Section 6.

Shifted Versions The next class of coupled systems is described by the approach of shifting the capacitive and resistive terms of the current coupling equations into the MNA yielding a new circuit with additional capacitors and possible resistors where the EM device is interpreted as a different mock element, see Lemma 5.8. Using any right-hand sides of (5.20) for (5.18) yields respectively the *shifted versions*

$$M_{\mathrm{E}} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_{\mathrm{E}}(\mathbf{u}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \mathbf{x}), \tag{5.24a}$$

$$f_{\mathrm{MNA2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA2}}(\mathbf{x}), \mathbf{x}, t) = c_{\mathrm{MNA2.1}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u}), \tag{5.24b}$$

$$M_{\mathrm{E}} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_{\mathrm{E}}(\mathbf{u}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}), \mathbf{x}), \tag{5.25a}$$

$$f_{\mathrm{MNA2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA2}}(\mathbf{x}), \mathbf{x}, t) = c_{\mathrm{MNA2.2}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u}). \tag{5.25b}$$

Note that such a shift is not possible for the choice $\mathbf{i}_\mathrm{E} = f_\mathrm{E0}(\mathbf{u})$ as there are no capacitive and resistive terms in the EM devices' constitutive equations. Moreover, the different topology in (5.24) and (5.25) guarantees, due to Lemma 5.7, the existence of capacitors parallel to the EM devices' branches which, especially, satisfies Assumption 3.16.

Remark 5.9
The coupled systems (5.24)-(5.25) *fulfills the Assumption 3.16 where EM devices represent mock elements in accordance with Lemma 5.8. Therefore, Corollary 3.17 applies these systems' MNA subsystems without any preconditions on the EM devices' positions.*

As shown in Chapter 6, this kind of subsystem arrangement does have an impact on waveform relaxation methods. Therefore, this choice of coupled system formulations may influence future interface design decisions.

Black-Box Versions Some EM modeling tools, such as the `devEM` modeler in [MAG16], provide interfaces that takes control of the current coupling themselves and only provide a voltage to current relation. That means both, the EM subsystem and the constitutive equations are in a so-called black box and therefore untouchable in their structure from outside. To mirror this structure, we are not allowed to incorporate the mock elements into the MNA model equations, cf. [MAG16; Sch+16]. Therefore, we call

$$M_\mathrm{E}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_\mathrm{E}(\mathbf{u}) = c_\mathrm{E}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}), \mathbf{x}), \tag{5.26a}$$

$$f_\mathrm{E0}(\mathbf{u}) = \mathbf{i}_\mathrm{E}, \tag{5.26b}$$

$$f_\mathrm{MNA1}(\frac{\mathrm{d}}{\mathrm{d}t}d_\mathrm{MNA1}(\mathbf{x}), \mathbf{x}, t) = c_\mathrm{MNA}(\mathbf{i}_\mathrm{E}), \tag{5.26c}$$

$$M_\mathrm{E}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_\mathrm{E}(\mathbf{u}) = c_\mathrm{E}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}), \mathbf{x}), \tag{5.27a}$$

$$f_\mathrm{E1}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{x}, \mathbf{u}) = \mathbf{i}_\mathrm{E}, \tag{5.27b}$$

$$f_\mathrm{MNA1}(\frac{\mathrm{d}}{\mathrm{d}t}d_\mathrm{MNA1}(\mathbf{x}), \mathbf{x}, t) = c_\mathrm{MNA}(\mathbf{i}_\mathrm{E}), \tag{5.27c}$$

$$M_\mathrm{E}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_\mathrm{E}(\mathbf{u}) = c_\mathrm{E}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}), \mathbf{x}), \tag{5.28a}$$

$$f_\mathrm{E2}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{x}, \mathbf{u}) = \mathbf{i}_\mathrm{E}, \tag{5.28b}$$

$$f_\mathrm{MNA1}(\frac{\mathrm{d}}{\mathrm{d}t}d_\mathrm{MNA1}(\mathbf{x}), \mathbf{x}, t) = c_\mathrm{MNA}(\mathbf{i}_\mathrm{E}) \tag{5.28c}$$

black-box versions. The system (5.26) was already intensively analyzed in [Bau12] and, therefore, we keep it as a reference. The other system formulations are justified as their current coupling equations are close to how they are implemented in `DevEM` though the real implementation is an industrial secret.

Proposition 5.10
All presented coupled systems, that is (5.21)-(5.28)*, are analytically equivalent in the following manner.*

(i) $(\mathbf{u}, \mathbf{x})$ *is a solution of either* (5.21)*,* (5.22)*,* (5.23)*,* (5.24) *and* (5.25) *or none of them.*

(ii) $(\mathbf{u}, \mathbf{x}, \mathbf{i}_\mathrm{E})$ *is a solution of either* (5.26)*,* (5.27) *and* (5.28) *or none of them.*

(iii) $(\mathbf{u}, \mathbf{x}, \mathbf{i}_\mathrm{E})$ *is a solution of* (5.26) *if and only if* $(\mathbf{u}, \mathbf{x})$ *is a solution of* (5.21) *and*

$$\mathbf{i}_\mathrm{E} = f_{\mathrm{E}0}(\mathbf{u}).$$

(iv) $(\mathbf{u}, \mathbf{x}, \mathbf{i}_\mathrm{E})$ *is a solution of* (5.27) *if and only if* $(\mathbf{u}, \mathbf{x})$ *is a solution of* (5.22) *and*

$$\mathbf{i}_\mathrm{E} = f_{\mathrm{E}1}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}, \mathbf{u}, \mathbf{x}).$$

(v) $(\mathbf{u}, \mathbf{x}, \mathbf{i}_\mathrm{E})$ *is a solution of* (5.28) *if and only if* $(\mathbf{u}, \mathbf{x})$ *is a solution of* (5.23) *and*

$$\mathbf{i}_\mathrm{E} = f_{\mathrm{E}2}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}, \mathbf{u}, \mathbf{x}).$$

Proof. The equivalence in (ii) follows by the equivalent construction of $f_{\mathrm{E}0}, f_{\mathrm{E}1}$ and $f_{\mathrm{E}2}$ while using MA (5.2b) and the canonical momentum (5.2c). For the same reasons are (5.21)-(5.23) equivalent and, moreover, (5.24) and (5.25). By definition of (5.16) and (5.17), the equivalence of (5.22) and (5.24) as well as (5.23) and (5.25) is obvious. From that follows (i). The equivalences in (iii)-(v) are also immediate consequences of the definitions. □

Remark 5.11
For all variants, the intrinsic equation of the mock element is the same, i. e. (5.21a)*,* (5.22a)*,* (5.23a)*,* (5.24a)*,* (5.25a)*,* (5.26a)*,* (5.27a) *and* (5.28a)*.*

Whereas, by Proposition 5.10, all coupled system variants are analytically equivalent, they might not be anymore when applying waveform relaxation schemes on them, see Section 6.3. In fact, concerning the incorporated and shifted versions, we see that even a little more structural information has already a huge effect on solving procedures.

5.3 Analysis of Coupled Field/Circuit DAEs

It is generally understood that when coupling two DAEs of a certain index each, the resulting monolithic DAE must not have an index equaling the maximum one nor of the accumulated sum, regardless of the index concept used. See for instance [Bau12] where the coupling of two DAEs with tractability index 1 and 2 may lead to an monolithic index-1 or index-2 DAE. Neither spared are the coupled field/circuit problems, as shown in Example 5.16. As a consequent it is not sufficient to use only index information about the subsystems in order to claim a certain index to the monolithic system.
In his treatise, Baumanns extended the well known topological index criterion for circuits, see e. g. [Tis99; ET00; Tis04], to the field/circuit coupled problem in the black-box variant (5.26) using the tractability index concept. Referring to this, a topological index criterion for the incorporated and shifted versions will be provided for the sake of completeness. Contrary to his analysis, the index concept used here is the previously introduce dissection index. Note that [Jan15] provides already a topological dissection index criterion for the MNA extended by various refined models but for a different EM device model then considered in Chapter 4. In the following we proof that the index statements of the new variants comply with those from Baumanns [Bau12] and, thus, there is no disadvantage from that to be expected when numerically solving these variants instead.
For a discussion on comparability of the two index concepts, e.g. that their indexes equal for linear systems, we refer to [Jan15].

All current coupling models (5.5), (5.6) and (5.7) for EM devices motivate the simplification introduced by

$$A_{\mathcal{E}} := A_{\mathrm{E}} \Lambda^{\top} G_{\Gamma}^{\top}.$$

The matrix $A_{\mathcal{E}}$ can be interpreted as an inflated incidence matrix of the EM devices by repetition of columns, insertions of zero columns and change of signs by Lemma 5.3.

The circuit coupled EM field device system (5.23) fits the form (2.2), that is,

$$f(\frac{\mathrm{d}}{\mathrm{d}t} d(x,t), x, t) = 0 \tag{5.29}$$

with $x = (\mathbf{e}, \mathbf{i}_{\mathrm{L}}, \mathbf{i}_{\mathrm{V}}, \boldsymbol{\phi}, \mathbf{a}, \boldsymbol{\pi})$, $w = (\mathbf{w}_{\mathrm{C}}, \mathbf{w}_{\mathrm{L}}, \mathbf{w}_{\mathrm{E}}, \mathbf{w}_{\phi}, \mathbf{w}_{\mathrm{a}}, \mathbf{w}_{\pi})$,

$$f(w,x,t) := \begin{pmatrix} A_{\mathrm{C}}\mathbf{w}_{\mathrm{C}} + A_{\mathcal{E}} M_{\varepsilon} \left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi} \right) + A_{\mathcal{E}}\mathbf{w}_{\mathrm{E}} + A_{\mathrm{R}} g_{\mathrm{R}}(A_{\mathrm{R}}^{\top}\mathbf{e}, t) + A_{\mathcal{E}} M_{\sigma} \left(G\boldsymbol{\phi} + A_{\mathcal{E}}^{\top}\mathbf{e} + \boldsymbol{\pi} \right) \\ + A_{\mathrm{L}}\mathbf{i}_{\mathrm{L}} + A_{\mathrm{V}}\mathbf{i}_{\mathrm{V}} + A_{\mathrm{I}} i_s(t) \\ \mathbf{w}_{\mathrm{L}} - A_{\mathrm{L}}^{\top}\mathbf{e} \\ v_s(t) - A_{\mathrm{V}}^{\top}\mathbf{e} \\ \widetilde{S} M_{\varepsilon} G \mathbf{w}_{\phi} + \widetilde{S} M_{\zeta} G M_{\xi} \widetilde{S} M_{\zeta} \mathbf{a} \\ \widetilde{C} M_{\nu} C \mathbf{a} + M_{\varepsilon} \left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi} \right) + \mathbf{w}_{\mathrm{E}} + M_{\sigma} \left(G\boldsymbol{\phi} + A_{\mathcal{E}}^{\top}\mathbf{e} + \mathbf{w}_{\mathrm{a}} \right) \\ \mathbf{w}_{\mathrm{a}} - \boldsymbol{\pi} \end{pmatrix}$$

and

$$d(x,t) := \begin{pmatrix} q_{\mathrm{C}}(A_{\mathrm{C}}^\top \mathbf{e}, t) \\ \phi_{\mathrm{L}}(\mathbf{i}_{\mathrm{L}}, t) \\ M_\varepsilon A_{\mathcal{E}}^\top \mathbf{e} \\ \boldsymbol{\phi} \\ \mathbf{a} \\ \boldsymbol{\pi} \end{pmatrix}.$$

Note that in case (5.29) is not an implicit ODE, it is a DAE with properly stated derivative term. In order to state any assessments concerning the dissection index of (5.29), we first provide an appropriate matrix chain as introduced in Section 2.3. The chain is similar to the one considered in [Jan15] and starts by defining

$$A := \begin{bmatrix} A_{\mathrm{C}} & 0 & A_{\mathcal{E}} & A_{\mathcal{E}} M_\varepsilon G & 0 & A_{\mathcal{E}} M_\varepsilon \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \widetilde{S} M_\varepsilon G & 0 & 0 \\ 0 & 0 & I & M_\varepsilon G & M_\sigma & M_\varepsilon \\ 0 & 0 & 0 & 0 & I & 0 \end{bmatrix},$$

$$D(x,t) := \begin{bmatrix} q_{\mathrm{C}}'(A_{\mathrm{C}}^\top \mathbf{e}, t) A_{\mathrm{C}}^\top & 0 & 0 & 0 & 0 & 0 \\ 0 & \phi_{\mathrm{L}}'(\mathbf{i}_{\mathrm{L}}, t) & 0 & 0 & 0 & 0 \\ M_\varepsilon A_{\mathcal{E}}^\top & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & I \end{bmatrix},$$

with

$$q_{\mathrm{C}}'(y,t) := \frac{\partial}{\partial y} q_{\mathrm{C}}(y,t) \quad \text{and} \quad \phi_{\mathrm{L}}'(y,t) := \frac{\partial}{\partial y} \phi_{\mathrm{L}}(y,t)$$

yielding

$$\begin{aligned} G_0(x,t) &:= AD(x,t) \\ &= \begin{bmatrix} A_{\mathrm{C}} q_{\mathrm{C}}'(A_{\mathrm{C}}^\top \mathbf{e}, t) A_{\mathrm{C}}^\top + A_{\mathcal{E}} M_\varepsilon A_{\mathcal{E}}^\top & 0 & 0 & A_{\mathcal{E}} M_\varepsilon G & 0 & A_{\mathcal{E}} M_\varepsilon \\ 0 & \phi_{\mathrm{L}}'(\mathbf{i}_{\mathrm{L}}, t) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \widetilde{S} M_\varepsilon G & 0 & 0 \\ M_\varepsilon A_{\mathcal{E}}^\top & 0 & 0 & M_\varepsilon G & M_\sigma & M_\varepsilon \\ 0 & 0 & 0 & 0 & I & 0 \end{bmatrix}. \end{aligned}$$

Further it is

$$B(x,t) := \begin{bmatrix} A_{\mathrm{R}} g_{\mathrm{R}}'(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\mathrm{R}}^\top + A_{\mathcal{E}} M_\sigma A_{\mathcal{E}}^\top & A_{\mathrm{L}} & A_{\mathrm{V}} & A_{\mathcal{E}} M_\sigma G & 0 & A_{\mathcal{E}} M_\sigma \\ -A_{\mathrm{L}}^\top & 0 & 0 & 0 & 0 & 0 \\ -A_{\mathrm{V}}^\top & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & H & 0 \\ M_\sigma A_{\mathcal{E}}^\top & 0 & 0 & M_\sigma G & \widetilde{C} M_\nu C & 0 \\ 0 & 0 & 0 & 0 & 0 & -I \end{bmatrix}$$

with

$$H := \widetilde{S} M_\zeta G M_\xi \widetilde{S} M_\zeta \quad \text{and} \quad g'_{\mathrm{R}}(y,t) := \frac{\partial}{\partial y} g_{\mathrm{R}}(y,t) \, .$$

Let P_{C} and Q_{C} be a kernel splitting pair associated with $A_{\mathrm{C}}^\top$, see Definition 2.11. We then introduce the notation

$$A_{\bar{\mathrm{C}}\mathrm{X}} := Q_{\mathrm{C}}^\top A_{\mathrm{X}} \, , \quad \text{for all} \quad \mathrm{X} \in \{\mathrm{V}, \mathrm{R}, \mathrm{L}, \mathrm{I}, \mathrm{M}\} \, .$$

The chain's first kernel splitting pairs $\{P, Q\}$ and $\{V, W\}$ of respectively AD and $(AD)^\top$ are introduced by

$$P := \begin{bmatrix} P_{\mathrm{C}} & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ -A_{\mathcal{E}}^\top P_{\mathrm{C}} & 0 & 0 & 0 & I \end{bmatrix} , \qquad V^\top := \begin{bmatrix} P_{\mathrm{C}}^\top & 0 & 0 & 0 & -P_{\mathrm{C}}^\top A_{\mathcal{E}} & P_{\mathrm{C}}^\top A_{\mathcal{E}} M_\sigma \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & I \end{bmatrix} ,$$

$$Q := \begin{bmatrix} Q_{\mathrm{C}} & 0 \\ 0 & 0 \\ 0 & I \\ 0 & 0 \\ 0 & 0 \\ -A_{\mathcal{E}}^\top Q_{\mathrm{C}} & 0 \end{bmatrix} , \qquad W^\top := \begin{bmatrix} Q_{\mathrm{C}}^\top & 0 & 0 & 0 & -Q_{\mathrm{C}}^\top A_{\mathcal{E}} & Q_{\mathrm{C}}^\top A_{\mathcal{E}} M_\sigma \\ 0 & 0 & I & 0 & 0 & 0 \end{bmatrix} .$$

Hence, we obtain

$$G_1(x,t) = V^\top G_0(x,t) P = \begin{bmatrix} P_{\mathrm{C}}^\top A_{\mathrm{C}} q'_{\mathrm{C}}(A_{\mathrm{C}}^\top \mathbf{e}, t) A_{\mathrm{C}}^\top & 0 & 0 & 0 & 0 & 0 \\ 0 & \phi'_{\mathrm{L}}(\mathbf{i}_{\mathrm{L}}, t) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \widetilde{S} M_\varepsilon G & 0 & 0 \\ M_\varepsilon A_{\mathcal{E}}^\top & 0 & 0 & M_\varepsilon G & M_\sigma & M_\varepsilon \\ 0 & 0 & 0 & 0 & I & 0 \end{bmatrix} P$$

$$= \begin{bmatrix} P_{\mathrm{C}}^\top A_{\mathrm{C}} q'_{\mathrm{C}}(A_{\mathrm{C}}^\top \mathbf{e}, t) A_{\mathrm{C}}^\top P_{\mathrm{C}} & 0 & 0 & 0 & 0 \\ 0 & \phi'_{\mathrm{L}}(\mathbf{i}_{\mathrm{L}}, t) & 0 & 0 & 0 \\ 0 & 0 & \widetilde{S} M_\varepsilon G & 0 & 0 \\ 0 & 0 & M_\varepsilon G & M_\sigma & M_\varepsilon \\ 0 & 0 & 0 & I & 0 \end{bmatrix} .$$

and, under consideration of Remark 2.15,

$$B_{y_1}^V(x,t) = V^\top B(x,t) Q$$

$$= \begin{bmatrix} P_{\mathrm{C}}^\top A_{\mathrm{R}} g'_{\mathrm{R}}(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\mathrm{R}}^\top & P_{\mathrm{C}}^\top A_{\mathrm{L}} & P_{\mathrm{C}}^\top A_{\mathrm{V}} & 0 & -P_{\mathrm{C}}^\top A_{\mathcal{E}} \widetilde{C} M_\nu C & 0 \\ -A_{\mathrm{L}}^\top & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & H & 0 \\ M_\sigma A_{\mathcal{E}}^\top & 0 & 0 & M_\sigma G & \widetilde{C} M_\nu C & 0 \\ 0 & 0 & 0 & 0 & 0 & -I \end{bmatrix} Q$$

$$= \begin{bmatrix} P_{\mathrm{C}}^\top A_{\mathrm{R}} g_{\mathrm{R}}'(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\mathrm{R}}^\top Q_{\mathrm{C}} & P_{\mathrm{C}}^\top A_{\mathrm{V}} \\ -A_{\mathrm{L}}^\top Q_{\mathrm{C}} & 0 \\ 0 & 0 \\ M_\sigma A_{\mathcal{E}}^\top Q_{\mathrm{C}} & 0 \\ A_{\mathcal{E}}^\top Q_{\mathrm{C}} & 0 \end{bmatrix} = \begin{bmatrix} P_{\mathrm{C}}^\top A_{\mathrm{R}} g_{\mathrm{R}}'(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\bar{\mathrm{C}}\mathrm{R}}^\top & P_{\mathrm{C}}^\top A_{\mathrm{V}} \\ -A_{\bar{\mathrm{C}}\mathrm{L}}^\top & 0 \\ 0 & 0 \\ M_\sigma A_{\bar{\mathrm{C}}\mathrm{M}}^\top & 0 \\ A_{\bar{\mathrm{C}}\mathrm{M}}^\top & 0 \end{bmatrix},$$

$$\begin{aligned} B_{x_1}^W(x,t) &= W^\top B(x,t) P \\ &= \begin{bmatrix} Q_{\mathrm{C}}^\top A_{\mathrm{R}} g_{\mathrm{R}}'(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\mathrm{R}}^\top & Q_{\mathrm{C}}^\top A_{\mathrm{L}} & Q_{\mathrm{C}}^\top A_{\mathrm{V}} & 0 & -Q_{\mathrm{C}}^\top A_{\mathcal{E}} \widetilde{C} M_\nu C & 0 \\ -A_{\mathrm{V}}^\top & 0 & 0 & 0 & 0 & 0 \end{bmatrix} P \\ &= \begin{bmatrix} Q_{\mathrm{C}}^\top A_{\mathrm{R}} g_{\mathrm{R}}'(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\mathrm{R}}^\top P_{\mathrm{C}} & Q_{\mathrm{C}}^\top A_{\mathrm{L}} & 0 & -Q_{\mathrm{C}}^\top A_{\mathcal{E}} \widetilde{C} M_\nu C & 0 \\ -A_{\mathrm{V}}^\top P_{\mathrm{C}} & 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} A_{\bar{\mathrm{C}}\mathrm{R}} g_{\mathrm{R}}'(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\mathrm{R}}^\top P_{\mathrm{C}} & A_{\bar{\mathrm{C}}\mathrm{L}} & 0 & -A_{\bar{\mathrm{C}}\mathrm{M}} \widetilde{C} M_\nu C & 0 \\ -A_{\mathrm{V}}^\top P_{\mathrm{C}} & 0 & 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} B_{y_1}^W(x,t) &= W^\top B(x,t) Q \\ &= \begin{bmatrix} Q_{\mathrm{C}}^\top A_{\mathrm{R}} g_{\mathrm{R}}'(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\mathrm{R}}^\top & Q_{\mathrm{C}}^\top A_{\mathrm{L}} & Q_{\mathrm{C}}^\top A_{\mathrm{V}} & 0 & -Q_{\mathrm{C}}^\top A_{\mathcal{E}} \widetilde{C} M_\nu C & 0 \\ -A_{\mathrm{V}}^\top & 0 & 0 & 0 & 0 & 0 \end{bmatrix} Q \\ &= \begin{bmatrix} Q_{\mathrm{C}}^\top A_{\mathrm{R}} g_{\mathrm{R}}'(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\mathrm{R}}^\top Q_{\mathrm{C}} & Q_{\mathrm{C}}^\top A_{\mathrm{V}} \\ -A_{\mathrm{V}}^\top Q_{\mathrm{C}} & 0 \end{bmatrix} = \begin{bmatrix} A_{\bar{\mathrm{C}}\mathrm{R}} g_{\mathrm{R}}'(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\bar{\mathrm{C}}\mathrm{R}}^\top & A_{\bar{\mathrm{C}}\mathrm{V}} \\ -A_{\bar{\mathrm{C}}\mathrm{V}}^\top & 0 \end{bmatrix}. \end{aligned}$$

We continue the chain by introducing a few more shorthands. Let $\{P_{\mathrm{V}}, Q_{\mathrm{V}}\}$ be a kernel splitting pair associated with $A_{\bar{\mathrm{C}}\mathrm{V}}^\top$ and $\{V_{\mathrm{V}}, W_{\mathrm{V}}\}$ the one for $A_{\bar{\mathrm{C}}\mathrm{V}}$. Following the above notation, we define

$$A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\mathrm{X}} := Q_{\mathrm{V}}^\top A_{\bar{\mathrm{C}}\mathrm{X}} = Q_{\mathrm{V}}^\top Q_{\mathrm{C}}^\top A_{\mathrm{X}}, \quad \text{for all} \quad \mathrm{X} \in \{\mathrm{R}, \mathrm{L}, \mathrm{I}, \mathrm{M}\}$$

and, with $\{P_{\mathrm{R}}, Q_{\mathrm{R}}\}$ being a kernel splitting pair of $A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\mathrm{R}}$,

$$A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{X}} := Q_{\mathrm{R}}^\top A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\mathrm{X}} = Q_{\mathrm{R}}^\top Q_{\mathrm{V}}^\top Q_{\mathrm{C}}^\top A_{\mathrm{X}}, \quad \text{for all} \quad \mathrm{X} \in \{\mathrm{L}, \mathrm{I}, \mathrm{M}\}.$$

The next basis functions for $B_{y_1}^W$ are then given by

$$Q_{y_1} = W_{y_1} = \begin{bmatrix} Q_{\mathrm{V}} Q_{\mathrm{R}} & 0 \\ 0 & W_{\mathrm{V}} \end{bmatrix}$$

yielding

$$\begin{aligned} Q_{y_1}^\top {B_{y_1}^V}^\top &= \begin{bmatrix} 0 & -Q_{\mathrm{R}}^\top Q_{\mathrm{V}}^\top Q_{\mathrm{C}}^\top A_{\mathrm{L}} & 0 & Q_{\mathrm{R}}^\top Q_{\mathrm{V}}^\top Q_{\mathrm{C}}^\top A_{\mathcal{E}} M_\sigma^\top & Q_{\mathrm{R}}^\top Q_{\mathrm{V}}^\top Q_{\mathrm{C}}^\top A_{\mathcal{E}} \\ W_{\mathrm{V}}^\top A_{\mathrm{V}}^\top P_{\mathrm{C}} & 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{L}} & 0 & A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{M}} M_\sigma^\top & A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{M}} \\ W_{\mathrm{V}}^\top A_{\mathrm{V}}^\top P_{\mathrm{C}} & 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

and

$$W_{y_1}^\top B_{x_1}^W = \begin{bmatrix} 0 & Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{L} & 0 & -Q_\mathrm{R}^\top Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathcal{E} \widetilde{C} M_\nu C & 0 \\ -W_\mathrm{V}^\top A_\mathrm{V}^\top P_\mathrm{C} & 0 & 0 & 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{L}} & 0 & -A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{M}} \widetilde{C} M_\nu C & 0 \\ -W_\mathrm{V}^\top A_\mathrm{V}^\top P_\mathrm{C} & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Further matrices are not calculated yet as we make use of the alternative ending described in Lemma 2.16. With this preliminaries we are now able to determine the dissection index of (5.23).

Theorem 5.12 (coupled system index)
Let Assumptions 3.7 and 3.8 be fulfilled. Further, the EM devices (E) meet Assumption 4.20 and their coupling satisfies Assumptions 5.1 and 5.2, each. Then, the coupled system (5.23)

(i) is an ODE if and only if there is a spanning tree in the circuit consisting only of capacitors and there are no voltage sources.

(ii) is a DAE having dissection index 1, if and only if it is not an ODE and there are no CV^+-loops and no LIE-cutsets.

(iii) is a DAE having at most dissection index 2.

Proof. The statements are proofed in ascending order.

(i) The topological conditions for the ODE case imply that $A_\mathrm{C}^\top$ has full column rank, according to Lemma A.5 and that A_V has no columns. Due to the latter, the third column and row block of $G_0(x,t)$ vanish, i. e.

$$G_0(x,t) = \begin{bmatrix} A_\mathrm{C} q_\mathrm{C}'(A_\mathrm{C}^\top \mathbf{e}, t) A_\mathrm{C}^\top + A_\mathcal{E} M_\varepsilon A_\mathcal{E}^\top & 0 & A_\mathcal{E} M_\varepsilon G & 0 & A_\mathcal{E} M_\varepsilon \\ 0 & \phi_\mathrm{L}'(\mathbf{i}_\mathrm{L}, t) & 0 & 0 & 0 \\ 0 & 0 & \widetilde{S} M_\varepsilon G & 0 & 0 \\ M_\varepsilon A_\mathcal{E}^\top & 0 & M_\varepsilon G & M_\sigma & M_\varepsilon \\ 0 & 0 & 0 & I & 0 \end{bmatrix}.$$

After left-multiplying and right-multiplying $G_0(x,t)$ with the nonsingular matrices

$$U^\top := \begin{bmatrix} I & 0 & 0 & -A_\mathcal{E}^\top & A_\mathcal{E}^\top M_\sigma \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix}, \qquad \widetilde{U} := \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ -A_\mathcal{E}^\top & 0 & 0 & 0 & I \end{bmatrix},$$

respectively, we obtain

$$\widetilde{G}_0(x,t) := U^\top G_0(x,t) \widetilde{U} = \begin{bmatrix} A_\mathrm{C} q_\mathrm{C}'(A_\mathrm{C}^\top \mathbf{e}, t) A_\mathrm{C}^\top & 0 & 0 & 0 & 0 \\ 0 & \phi_\mathrm{L}'(\mathbf{i}_\mathrm{L}, t) & 0 & 0 & 0 \\ 0 & 0 & \widetilde{S} M_\varepsilon G & 0 & 0 \\ 0 & 0 & M_\varepsilon G & M_\sigma & M_\varepsilon \\ 0 & 0 & 0 & I & 0 \end{bmatrix}.$$

Due to the full column rank of $A_{\mathrm{C}}^\top$, the positive definitness of $q'_{\mathrm{C}}(A_{\mathrm{C}}^\top \mathbf{e}, t)$ and $\phi'_{\mathrm{L}}(\mathbf{i}_{\mathrm{L}}, t)$ and due to the nonsingularity of the 3×3-block, see Lemma 5.4, belonging to the electromagnetic quantities, $\widetilde{G}_0(x,t)$ is nonsingular and thus is $G_0(x,t)$.

(ii) Respectively, the topological conditions for the index-1 case are equivalent to $Q_{\mathrm{C}}^\top A_{\mathrm{V}}$ having full column rank, see Lemma A.7, and $\begin{bmatrix} A_{\mathrm{C}} & A_{\mathrm{R}} & A_{\mathrm{V}} \end{bmatrix}$ having full row rank, according to Lemma A.6. Further, due to Lemma A.8, the latter condition is equivalent to $Q_{\mathrm{V}}^\top Q_{\mathrm{C}}^\top A_{\mathrm{R}}$ having full row rank. Let $(u, v) \in \ker B_{y_1}^W$, that is

$$\begin{aligned} Q_{\mathrm{C}}^\top A_{\mathrm{R}} g'_{\mathrm{R}}(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\mathrm{R}}^\top Q_{\mathrm{C}} u + Q_{\mathrm{C}}^\top A_{\mathrm{V}} v &= 0, \\ -A_{\mathrm{V}}^\top Q_{\mathrm{C}} u &= 0. \end{aligned}$$

From the second equation follows $u \in \ker(A_{\mathrm{V}}^\top Q_{\mathrm{C}})$ meaning that there exists a z so that $u = Q_{\mathrm{V}} z$. The first equation hence yields, after left-multiplying with $Q_{\mathrm{V}}^\top$

$$Q_{\mathrm{V}}^\top Q_{\mathrm{C}}^\top A_{\mathrm{R}} g'_{\mathrm{R}}(A_{\mathrm{R}}^\top \mathbf{e}, t) A_{\mathrm{R}}^\top Q_{\mathrm{C}} Q_{\mathrm{V}} z = 0 \quad \Leftrightarrow \quad z = 0$$

as $g'_{\mathrm{R}}(A_{\mathrm{R}}^\top \mathbf{e}, t)$ is assumed to be positive definite and $Q_{\mathrm{V}}^\top Q_{\mathrm{C}}^\top A_{\mathrm{R}}$ has full row rank. Therefore, $u = 0$.
Since $Q_{\mathrm{C}}^\top A_{\mathrm{V}}$ has full column rank it follow from $Q_{\mathrm{C}}^\top A_{\mathrm{V}} v = 0$ that $v = 0$.
We conclude that $B_{y_1}^W$ is nonsingular yielding the index-1 condition.

(iii) In order to check that the index does not exceeded 2, we have to prove that ${W_y^*}^\top G_1(x,t) Q_{x_1}$ is nonsingular, see Lemma 2.16, where W_y^* and Q_{x_1} are the basis (matrix) functions of the kernels associated with $Q_{y_1}^\top {B_{y_1}^V}^\top$ and $W_{y_1}^\top B_{x_1}^W$, respectively. This condition is equivalent to

$${W_y^*}^\top G_1(x,t) Q_{x_1} z = 0 \quad \Rightarrow \quad z = 0\,.$$

As $Q_{x_1} z \in \operatorname{im} Q_{x_1}$ it is by construction also in $\ker W_{y_1}^\top B_{x_1}^W$. Hence, we have to show that from

$${W_y^*}^\top G_1(x,t) u = 0 \quad \text{and} \quad W_{y_1}^\top B_{x_1}^W u = 0 \quad \Rightarrow \quad u = 0.$$

Since $\ker {W_y^*}^\top = \operatorname{im} B_{y_1}^V Q_{y_1}$ and $\operatorname{im} Q_{x_1} = \ker W_{y_1}^\top B_{x_1}^W$, the latter statement is equivalent to

$$G_1(x,t) u \in \operatorname{im} B_{y_1}^V Q_{y_1} \quad \text{and} \quad W_{y_1}^\top B_{x_1}^W u = 0 \quad \Rightarrow \quad u = 0 \tag{5.30}$$

which we are going to show in the following:

- Let $G_1(x,t) u \in \operatorname{im} B_{y_1}^V Q_{y_1}$ with $u = (u_1, u_2, u_3, u_4, u_5)$, then there exist (y_1, y_2) such that

$$\begin{aligned} P_{\mathrm{C}}^\top A_{\mathrm{C}} q'_{\mathrm{C}}(A_{\mathrm{C}}^\top \mathbf{e}, t) A_{\mathrm{C}}^\top P_{\mathrm{C}} u_1 &= P_{\mathrm{C}}^\top A_{\mathrm{V}} W_{\mathrm{V}} y_1\,, && (5.31) \\ L(\mathbf{i}_{\mathrm{L}}, t) u_2 &= -A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{L}}^\top y_2\,, && (5.32) \\ \widetilde{S} M_\varepsilon G u_3 &= 0\,, && (5.33) \\ M_\varepsilon G u_3 + M_\sigma u_4 + M_\varepsilon u_5 &= M_\sigma A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{M}}^\top y_2\,, && (5.34) \\ u_4 &= A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{M}}^\top y_2 && (5.35) \end{aligned}$$

and assume that further $W_{y_1}^\top B_{x_1}^W u = 0$, i. .e.

$$A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{L}} u_2 - A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{M}} \widetilde{C} M_\nu C u_4 = 0 \,, \tag{5.36}$$

$$-W_{\mathrm{V}}^\top A_{\mathrm{V}}^\top P_{\mathrm{C}} u_1 = 0 \,. \tag{5.37}$$

Multiplying (5.31) with $y_1^\top$ from left and (5.37) with u_1 from right yields

$$y_1^\top P_{\mathrm{C}}^\top A_{\mathrm{C}} q_{\mathrm{C}}'(A_{\mathrm{C}}^\top \mathbf{e}, t) A_{\mathrm{C}}^\top P_{\mathrm{C}} u_1 = y_1^\top P_{\mathrm{C}}^\top A_{\mathrm{V}} W_{\mathrm{V}} u_1 = 0 \,.$$

Since $q_{\mathrm{C}}'(A_{\mathrm{C}}^\top \mathbf{e}, t)$ is positive definite, by assumption, and $A_{\mathrm{C}}^\top P_{\mathrm{C}}$ has full col rank, by construction, we obtain $u_1 = 0$.

- Because of the full row rank property of $\widetilde{S} = -G^\top$ and the positive definitness of M_ε, see Assumption 4.20, it follows from (5.33) that $u_3 = 0$.
- Furthermore, it is also $\phi_{\mathrm{L}}'(\mathbf{i}_{\mathrm{L}}, t)$ positive definite by assumption, guaranteeing the existence of ${\phi_{\mathrm{L}}'}^{-1}(\mathbf{i}_{\mathrm{L}}, t)$. Multiplying (5.32) from left with ${\phi_{\mathrm{L}}'}^{-1}$ and inserting u_2 and u_4 from (5.35) into (5.36) yields

$$-A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{L}} {\phi_{\mathrm{L}}'}^{-1}(\mathbf{i}_{\mathrm{L}}, t) A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{L}}^\top y_2 - A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{M}} \widetilde{C} M_\nu C A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{M}}^\top y_2 = 0 \,. \tag{5.38}$$

  Again, we know that ${\phi_{\mathrm{L}}'}^{-1}(\mathbf{i}_{\mathrm{L}}, t)$ has to be positive definite and the same holds for $\widetilde{C} M_\nu C$, due to the assumptions for M_ν and since $\widetilde{C} = C^\top$. Hence, from (5.38) we deduce that $y_2 = 0$ if $\begin{bmatrix} A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{L}}^\top \\ A_{\bar{\mathrm{C}}\bar{\mathrm{V}}\bar{\mathrm{R}}\mathrm{M}}^\top \end{bmatrix}$ has full column rank. Lemma A.8 implies now that this holds true if and only if there is no I-cutset. As of Assumption 3.8 we do not allow I-cutsets leading to $y_2 = 0$. It immediately follows from (5.32), due to the nonsingularity of $\phi_{\mathrm{L}}'(\mathbf{i}_{\mathrm{L}}, t)$, that $u_2 = 0$.
- From (5.35) that $u_4 = 0$.
- Finally, from (5.34), after exploiting $u_3 = 0$, $u_4 = 0$ and the nonsingularity of M_ε, that $u_5 = 0$.

In conclusion, we figure that $u = 0$ and statement (5.30) is valid. Hence, index-2 is not exceeded.

□

Lemma 5.13 (exchanging $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{a}$ and $\boldsymbol{\pi}$)
Having the quasi-canonical momentum $\boldsymbol{\pi}$ replaced by the magnetic vector potential's derivative $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{a}$ and/or vice versa in either the coupling equation or Maxwell-Ampère's law of (5.29) does not change the dissection index chain but for the first set of kernel splitting pairs.

Proof. Going through all the possible permutations, the chain's first set of kernel splitting pairs for

$$\bar{f}(w,x,t) := \begin{pmatrix} A_{\mathrm{C}}\mathbf{w}_{\mathrm{C}} + A_{\mathcal{E}}M_{\varepsilon}\left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi}\right) + A_{\mathcal{E}}\mathbf{w}_{\mathrm{E}} + A_{\mathrm{R}}g_{\mathrm{R}}(A_{\mathrm{R}}^{\top}\mathbf{e},t) + A_{\mathcal{E}}M_{\sigma}\left(G\boldsymbol{\phi} + A_{\mathcal{E}}^{\top}\mathbf{e} + \mathbf{w}_{\mathrm{a}}\right) \\ +A_{\mathrm{L}}\mathbf{i}_{\mathrm{L}} + A_{\mathrm{V}}\mathbf{i}_{\mathrm{V}} + A_{\mathrm{I}}i_s(t) \\ \mathbf{w}_{\mathrm{L}} - A_{\mathrm{L}}^{\top}\mathbf{e} \\ v_s(t) - A_{\mathrm{V}}^{\top}\mathbf{e} \\ \widetilde{S}M_{\varepsilon}G\mathbf{w}_{\phi} + \widetilde{S}M_{\zeta}GM_{\xi}\widetilde{S}M_{\zeta}\mathbf{a} \\ \widetilde{C}M_{\nu}C\mathbf{a} + M_{\varepsilon}\left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi}\right) + \mathbf{w}_{\mathrm{E}} + M_{\sigma}\left(G\boldsymbol{\phi} + A_{\mathcal{E}}^{\top}\mathbf{e} + \mathbf{w}_{\mathrm{a}}\right) \\ \mathbf{w}_{\mathrm{a}} - \boldsymbol{\pi} \end{pmatrix}$$

or

$$\bar{\bar{f}}(w,x,t) := \begin{pmatrix} A_{\mathrm{C}}\mathbf{w}_{\mathrm{C}} + A_{\mathcal{E}}M_{\varepsilon}\left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi}\right) + A_{\mathcal{E}}\mathbf{w}_{\mathrm{E}} + A_{\mathrm{R}}g_{\mathrm{R}}(A_{\mathrm{R}}^{\top}\mathbf{e},t) + A_{\mathcal{E}}M_{\sigma}\left(G\boldsymbol{\phi} + A_{\mathcal{E}}^{\top}\mathbf{e} + \boldsymbol{\pi}\right) \\ +A_{\mathrm{L}}\mathbf{i}_{\mathrm{L}} + A_{\mathrm{V}}\mathbf{i}_{\mathrm{V}} + A_{\mathrm{I}}i_s(t) \\ \mathbf{w}_{\mathrm{L}} - A_{\mathrm{L}}^{\top}\mathbf{e} \\ v_s(t) - A_{\mathrm{V}}^{\top}\mathbf{e} \\ \widetilde{S}M_{\varepsilon}G\mathbf{w}_{\phi} + \widetilde{S}M_{\zeta}GM_{\xi}\widetilde{S}M_{\zeta}\mathbf{a} \\ \widetilde{C}M_{\nu}C\mathbf{a} + M_{\varepsilon}\left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi}\right) + \mathbf{w}_{\mathrm{E}} + M_{\sigma}\left(G\boldsymbol{\phi} + A_{\mathcal{E}}^{\top}\mathbf{e} + \boldsymbol{\pi}\right) \\ \mathbf{w}_{\mathrm{a}} - \boldsymbol{\pi} \end{pmatrix}$$

instead of f reads

$$P := \begin{bmatrix} P_{\mathrm{C}} & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ -A_{\mathcal{E}}^{\top}P_{\mathrm{C}} & 0 & 0 & 0 & I \end{bmatrix}, \qquad V^{\top} := \begin{bmatrix} P_{\mathrm{C}}^{\top} & 0 & 0 & 0 & -P_{\mathrm{C}}^{\top}A_{\mathcal{E}} & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & I \end{bmatrix},$$

$$Q := \begin{bmatrix} Q_{\mathrm{C}} & 0 \\ 0 & 0 \\ 0 & I \\ 0 & 0 \\ 0 & 0 \\ -A_{\mathcal{E}}^{\top}Q_{\mathrm{C}} & 0 \end{bmatrix}, \qquad W^{\top} := \begin{bmatrix} Q_{\mathrm{C}}^{\top} & 0 & 0 & 0 & -Q_{\mathrm{C}}^{\top}A_{\mathcal{E}} & 0 \\ 0 & 0 & I & 0 & 0 & 0 \end{bmatrix}.$$

Making use of

$$\bar{\bar{\bar{f}}}(w,x,t) :=$$

$$\begin{pmatrix} A_{\mathrm{C}}\mathbf{w}_{\mathrm{C}} + A_{\mathcal{E}}M_{\varepsilon}\left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi}\right) + A_{\mathcal{E}}\mathbf{w}_{\mathrm{E}} + A_{\mathrm{R}}g_{\mathrm{R}}(A_{\mathrm{R}}^{\top}\mathbf{e}, t) + A_{\mathcal{E}}M_{\sigma}\left(G\boldsymbol{\phi} + A_{\mathcal{E}}^{\top}\mathbf{e} + \mathbf{w}_{\mathrm{a}}\right) \\ +A_{\mathrm{L}}\mathbf{i}_{\mathrm{L}} + A_{\mathrm{V}}\mathbf{i}_{\mathrm{V}} + A_{\mathrm{I}}i_s(t) \\ \mathbf{w}_{\mathrm{L}} - A_{\mathrm{L}}^{\top}\mathbf{e} \\ v_s(t) - A_{\mathrm{V}}^{\top}\mathbf{e} \\ \widetilde{S}M_{\varepsilon}G\mathbf{w}_{\phi} + \widetilde{S}M_{\zeta}GM_{\xi}\widetilde{S}M_{\zeta}\mathbf{a} \\ \widetilde{C}M_{\nu}C\mathbf{a} + M_{\varepsilon}\left(G\mathbf{w}_{\phi} + \mathbf{w}_{\pi}\right) + \mathbf{w}_{\mathrm{E}} + M_{\sigma}\left(G\boldsymbol{\phi} + A_{\mathcal{E}}^{\top}\mathbf{e} + \boldsymbol{\pi}\right) \\ \mathbf{w}_{\mathrm{a}} - \boldsymbol{\pi} \end{pmatrix}$$

instead of f lets the set read

$$P := \begin{bmatrix} P_{\mathrm{C}} & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ -A_{\mathcal{E}}^{\top}P_{\mathrm{C}} & 0 & 0 & 0 & I \end{bmatrix}, \qquad V^{\top} := \begin{bmatrix} P_{\mathrm{C}}^{\top} & 0 & 0 & 0 & -P_{\mathrm{C}}^{\top}A_{\mathcal{E}} & -P_{\mathrm{C}}^{\top}A_{\mathcal{E}}M_{\sigma} \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & I \end{bmatrix},$$

$$Q := \begin{bmatrix} Q_{\mathrm{C}} & 0 \\ 0 & 0 \\ 0 & I \\ 0 & 0 \\ 0 & 0 \\ -A_{\mathcal{E}}^{\top}Q_{\mathrm{C}} & 0 \end{bmatrix}, \qquad W^{\top} := \begin{bmatrix} Q_{\mathrm{C}}^{\top} & 0 & 0 & 0 & -Q_{\mathrm{C}}^{\top}A_{\mathcal{E}} & -Q_{\mathrm{C}}^{\top}A_{\mathcal{E}}M_{\sigma} \\ 0 & 0 & I & 0 & 0 & 0 \end{bmatrix}.$$

Simple multiplication shows that the forthcoming chain members, which are $G_1, B_{y_1}^V, B_{x_1}^W, B_{y_1}^W$ etc., coincide with those of f. □

Corollary 5.14
The statement of Theorem 5.12 also applies further to the following coupled systems:

(i) the incorporated version (5.22),

(ii) the shifted version (5.24),

(iii) the shifted version (5.25).

Proof. The equation (5.22) is given for $\bar{\bar{f}}(w, x, t) = 0$ with $\bar{f}$ of Lemma 5.13. Exploiting the additive coupled system structure of (5.24) and (5.25), their Jacobians comply respectively with those of (5.22) and (5.23). □

Concerning the topological interpretation of the incorporated and their corresponding shifted versions, see for instance Figure 5.3, one has to be careful in the following way:

Remark 5.15
The shifted interpretation of an EM device as capacitors, probable resistors and a mock element, according to Lemma 5.8, does not impose entire topologically independent capacitors which affect the CV^{+}-loop or LIE-cutset property of the topological index criterion in Theorem 5.12. Their influence as such is canceled out by the remaining mock element part.

The following example points out the crucial fact that the index of a DAE, representing a coupled system, may differ from its individual subsystems' indices. It further supports the statement of Remark 5.15.

Example 5.16
Consider the coupled field/circuit given in Figure 5.4 using the formulation in (5.24). The

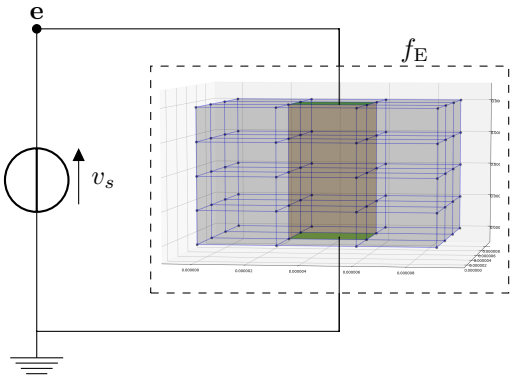


Figure 5.4: Coupled circuit with a voltage source and an EM device acting as an 100 Ohm resistor.

resulting system has index 1, as of Theorem 5.12, and reads

$$-\mathbf{i}_{\mathrm{V}} + q_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{e}) + g_{\mathrm{E}}(\mathbf{e}) = -s_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u}), \tag{5.39a}$$

$$v_s(t) - \mathbf{e} = 0, \tag{5.39b}$$

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_{\mathrm{E}}(\mathbf{u}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{e}, \mathbf{e}) \tag{5.39c}$$

where $\mathbf{i}_{\mathrm{V}}$ and $\mathbf{e}$ are now scalar valued. Splitting (5.39) into two subsystems according to Kirchhoff's laws and the EM equations, i. e. (5.39a)-(5.39b) as the first and (5.39c) as the second, we treat the variables in $\mathbf{u}$ and $(\mathbf{e}, \mathbf{i}_{\mathrm{V}})$ as parameters for the first and second subsystem, respectively. Hence, we can write down the first subsystem as

$$\begin{aligned} -\mathbf{i}_{\mathrm{V}} + q_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{e}) + g_{\mathrm{E}}(\mathbf{e}) &= r(t), \\ v_s(t) - \mathbf{e} &= 0, \end{aligned}$$

for some given right-hand side r, and observe that it has index 2 since q_{E} is positive definite by assumption. As a matter of fact, shifting the circuit variable influences of the current coupling (5.16) or (5.17) into the MNA, similar as in Figure 5.3, and erasing the mock element part, leaves a topological circuit as in Figure 5.5 where the resistors might be absence in case $g_{\mathrm{E}} = 0$.

The index-2 property does not occur surprisingly since the first subsystem's circuit has indeed a CV-loop and hence complying with the standard topological index property of [ET00].

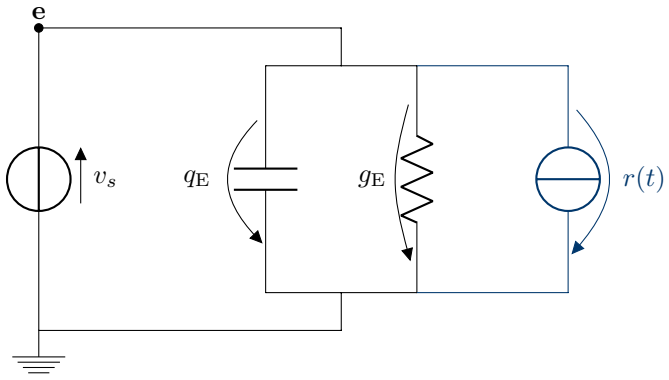


Figure 5.5: Circuit's subsystem (5.39a)-(5.39b) of a shifted variant with capacitors and resistors shifted from the EM device's coupling equation into the MNA. The coupling equation's residual part, the right-hand side r, represented as a current source.

Conclusion

In this chapter we provided two possibilities of how to incorporate the EM field device model, introduced in Chapter 4, as a refined model into a circuit, modeled by the MNA, using the mock element introduced in Chapter 3. Either the EM field device model can be interpreted directly as a mock element, see Lemma 5.6, or it can be interpreted as an equivalent circuit consisting of parallel branches of capacitors, possible resistors and another mock element, see Lemma 5.6.

We have seen that there are three variations of the current coupling equation, linking the EM device model with the electric circuit model. These are (5.10) as used in [Bau12], as well as (5.11) and (5.12), which are likely to be used by the device electromagnetic modeler `DevEM` from *MAGWEL*, for which we originally designed the circuit interface via the mock element. The exact coupling or model equations are an industrial secret. From these three possibilities and two mock element interpretations, we derived eight analytically equivalent field/circuit coupled systems (5.21)-(5.28). For some representative systems (5.22)-(5.25) we derived a topological index result using the dissection index concept, see Theorem 5.12 and Corollary 5.14. Note that, the coupling variants involving the introduction of additional capacitors do not influence the topological criteria of Theorem 5.12. The index results coincide with those from [Bau12], where on the contrary the tractability index is used. They also comply with the statement in [Jan15], where a different EM device model is used. Last but not least, the index analysis agrees on the results in [Cor+20], where the here considered refined model fits into the concept of generalized elements as an inductance-like one. All of these statements can be considered as generalizations of the topological criteria in [Tis99]. We finished this chapter with Example 5.16 emphasizing that we cannot anticipate a coupled systems index only based upon the subsystems individual indices.

6 Simulation of Field/Circuit Coupled Systems

When simulating physical phenomena, we usually develop models that compromise between a bunch of criteria such as importance, accuracy, knowledge, avoiding complexity etc. For example, electric circuit modeling in Section 3 uses modified nodal analysis (MNA) and electromagnetic (EM) devices in Section 4 are modeled using Maxwell's equations (MEs) and finite-integration-technique (FIT). Therefore, problems involving multiple physical phenomena are typically represented by collections of these individual models coupled with each other, for instance the circuit coupled EM device problem of Section 5 as a combination of modeling approaches in Chapter 3 and 4, resulting in coupled systems of differential-algebraic equations (DAEs), see e.g. [LBH84; Tis04].

Depending on how strong these models subsystems are coupled within the overall problem's system, one can think of different approaches to numerically solve them. One possibility could be to solve the whole system in a *monolithic* way, meaning all subsystems combined as one, see e. g. [Ben06; Sch+16; Sch+19b]. Another approach could be to address the subsystems individually in an either sequential or parallel way, which we refer to as *co-simulation*, see e. g. [Lel82; WS12; ST20]. With regard to the latter approach, a further distinction is made between various coupling schemes, see Figure 6.1. Whereas, for the monolithic approach, all

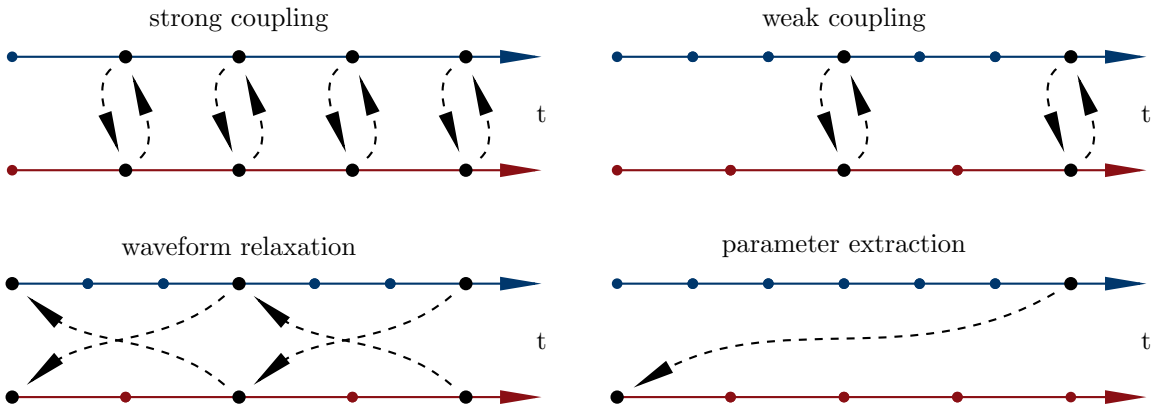


Figure 6.1: Different coupling schemes based on [Sch11; Cle+12].

subsystems have to be solved for the same time points by the same numerical method, the co-simulation approach allows for solving the subsystems on an individual time scale with possible tailored numerical methods, that are those which might exploit special structural properties. For example, the highly desirable leapfrog integration scheme for the EM device's subsystem,

if given $\mathbf{E}-\mathbf{H}$ formulation, see [Yee66; Wei96; Bos99; SW01]. Thus, one can profit of the integrator's symplectic and explicit properties. On the other hand, with the co-simulation approach we require an additional orchestration of these individual processes, cf. [Pus+14].

In order to solve a multi-physical problem, one can think of combining different industrial tools, which are each superior in its field. Building a simulator for combination thereof, leads to difficulties as the industrial modelers are usually black boxes, see e. g. [Kan01; Zho+06]. In other words, we most likely do not have access to the model equations or underlying DAE functions but we can get solutions of the individual subsystems buy running the tools' solver. In this way we can solve the coupled problem with the co-simulation approach, but not monolithically. To circumvent this problem and enable a monolithic simulation, a suitable interface is required through which the tools can communicate.

Motivated by the collaboration with *MAGWEL* within the *nanoCOPS* project, we present a nonlinear interface for their EM device modeler `DevEM` which allows to communicate with circuit simulators of other parties. The design of this interface is based on that of the mock element introduced in Chapter 3, with which the coupled field/circuit modeling in Chapter 5 is archived.
In this chapter, we present the coupled field/circuit modeling and simulation framework, see further [Sch+16; Bit+18; Mat+19]. The in-house electric circuit simulator's design, which we used to test the interface with `DevEM`, is based on those from the semiconductor industry and makes use of lumped elements described by netlists. This enabled us to better understand the demands on the interface from an industrial perspective.
In Section 6.2, we show that this framework allows for a monolithic simulation approach, contrary to [SDB10; SDB12] where a co-simulation approach is presented. We monolithically solve the coupled systems given in Chapter 5 using time integration methods for DAEs, see e. g. [HW91; LMT13]. Hereby, a large part of the challenge is represented by solving the underlying linear systems of equations, especially when dealing with huge dimensions. In order to circumvent these problems, we introduce various methods such as scaling methods, discuss direct and iterative solvers, propose a hybrid method for solvers and make use of structural exploitation. As the refined model equations from third party software are usually black-box, we use the model equations presented in Chapter 5. As a by-product of the monolithic simulation possibility, it is also possible to use co-simulation approaches. Therefore, in Section 6.3 we focus on *waveform relaxation* (WR) methods, sometimes called *dynamic iteration*, which is a frequently used candidate for the co-simulation approach concerning field/circuit coupled systems, see e. g. [BBS14; AG01; SDB10]. Since convergence of waveform relaxation methods is not always guaranteed, as soon as DAEs are involved, see e. g. [AG01], we provide some topological criteria addressing that topic in Section 6.3.3, see also [PT18; ST20]. These results are numerically supported by benchmarks in Chapter 7 using a self-implemented version of the here provided field/circuit models and the refined models provided by `DevEM` for some test cases.

The chapter is structured as follows. First we introduce the interface concept to combine circuit and EM field devices. Secondly, we focus on monolithic solving methods for the field/circuit coupled systems of Chapter 5 including DAE time integration methods, linear solving strategies, scaling and structural exploitation. Then, we introduce waveform relaxation schemes to these

coupled systems with a focus on Gauss-Seidel and Jacobi types. We conclude this chapter with a convergence analysis upon the preparatory work in Chapter 3 yielding topological criteria for these WR schemes.

6.1 Interface Realization

In order to incorporate spatially distributed equations from the DevEM into a circuit modeler, an interface is required. If the model equations can be represented by matrices only, an interface can be realized by simply exchanging them. Consequently, these refined model equations have to be entirely linear. This obstacle can be circumvented by a nonlinear interface which hence also allows for semi-conducting material to be considered. Contrary to the matrix exchange this nonlinear interface provides point-wise function evaluations of the refined model equations.

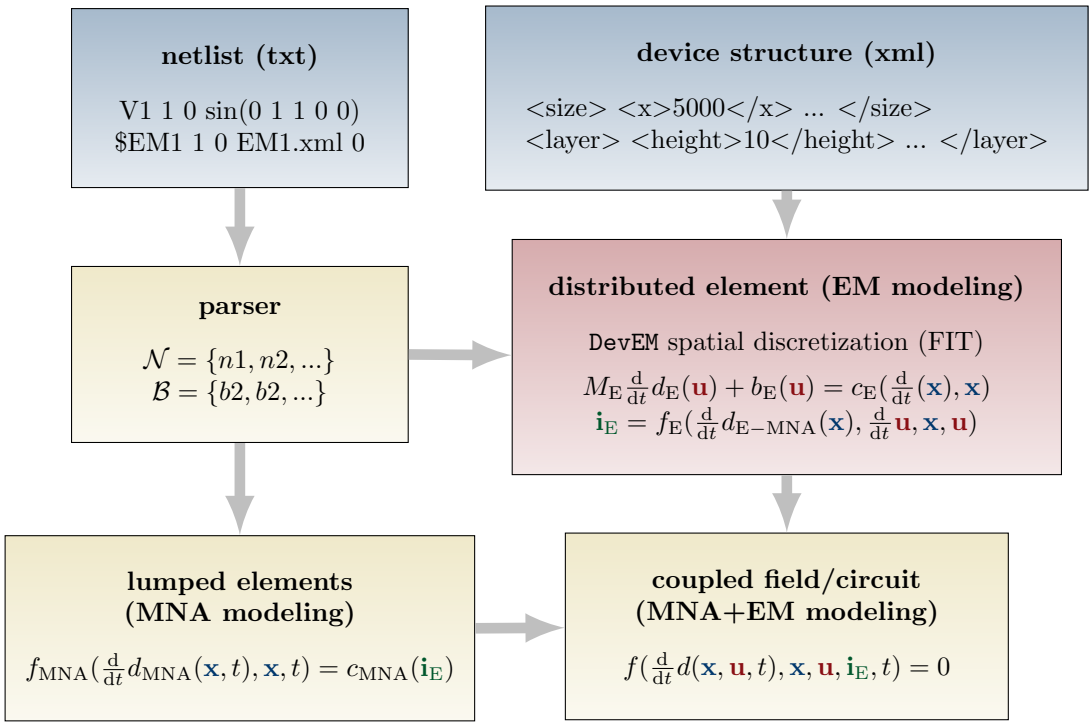


Figure 6.2: Coupled field/circuit modeling framework for DevEM.

With this interface set up, the monolithic modeling is archived via the framework given in Figure 6.2. In particular, for the in-house circuit simulator and DevEM, the monolithic modeling using the nonlinear interface is performed as follows:

(i) The circuit structure and lumped elements are stored in so-called netlists. For EM devices the according line begins with $EM followed by the topological position, i. e. the nodes that are connected to the EM device's contacts, and the EM device's structure filename.

(ii) The circuit simulator's parser processes the netlist file and initializes the graph structure. When triggered by $EM, it calls the EM device modeler, for instance `DevEM` with the according structure filename.

(iii) The corresponding structure file contains the EM device's geometric and material structure in terms of an xml format. It is interpreted by the device modeler which generates the refined model equations, here using FIT for the spatial discretization. Via the nonlinear interface, evaluations of these equations and their derivatives, together with some meta data such as dimensional information, are provided.

(iv) The circuit simulator initializes the equations for the lumped elements and takes into account the additional quantities caused by the refined models.

(v) For the simulation process, the circuit simulator orchestrates all the equations' evaluations, including the ones accessed by the interface, mimicking the desired formulation of the coupled model. The resulting DAE description can then be processed by an appropriate solver.

Note that contrary to the mock element introduced in Chapter 3, the EM device's intrinsic equations obtained by `DevEM` must not represent an ODE. As already discussed, we may expect a DAE instead of an ODE.

6.2 Monolithically Solving the Field/Circuit Coupled Problem

Coupled circuit and EM device problems usually lead to model equations with unknowns easily in the millions. It is fair enough to say that the large number's origin is due to the EM device's refined model equations, especially if there is more then just one EM device to be considered in the coupled framework. Surely, the research community undergoes efforts to provide methods for calculating economical computational meshes that are still capable of meeting industrial requirements, see e. g. [PCW03; Car97]. Also simplifications of the electromagnetic field models such as quasistatic and static models can be considered. However, during time integration one ends up solving a linear system of equations which can be addressed in a similar way as solving the saddle point problems, see Section 6.2.3 and furthermore [Sch11; Cle+11; BGL05]. But not only with focus on the subsystems there exist approaches, for instance one can bypass the field subproblem during time integration with the help of Schur complement and only update it if possible, see e. g. [SBD12]. Others improve the solving process of field/circuit coupled systems with the help of *model order reduction* methods, see e. g. [SVR08; BF15]. If there are fast and slow scale signals present, relative to each other, so-called *multirate PDAE* ansatz can be applied, see e. g. [Bra+96; Bra01; PG02], as realized withing the *nanoCOPS* project [Bit+18]. Doing so one should always keep in mind next step of numerically solving the resulting systems and find a satisfying compromise which is yet an open subject to research, see e. g. [Bar+16; Ben06]. In case that the modeling routines are black boxes, there are not much possibilities left. After a general introduction of time integration and rootfinding methods, we study techniques such as hybrid linear solvers, scaling and structural exploitation which can be realized without interfering to much with the modeling.

6.2.1 Time Integration

The first step in solving a DAEs numerically is usually by means of time integration methods. Here we make use of well established ODE time integration methods though not all of them are suited for that purpose [HW91]. For instance, explicit methods can not be applied out of the box, leading to investigation of so-called half-explicit methods [HLR89]. In the following we introduce some popular classes of methods that serve this purpose.

For a compact time interval $[t_0, T] = \mathcal{I} \subset \mathbb{R}$ with consistent initial condition $x(t_0) = x_0$, consider the DAE

$$f(\frac{\mathrm{d}}{\mathrm{d}t} d(x(t), t), x(t), t) = 0. \tag{6.1}$$

Let further $t_n \in \mathcal{I}$, for $n \in \{1, \ldots, N\}$, be sufficiently many interpolation points satisfying $t_0 < t_1 < \cdots < t_N = T$. With $h_n := t_n - t_{n-1}$ we denote the step size and the numerical approximations at time point t_n are stored in x_n.

Backward Differentiation Formulas As an implicit linear multi-step method, the *backward differentiation formula* (BDF) is a widely approved method for solving DAEs, see e. g. [Gea71]. In case of a DAE with nonlinear derivative term (6.1), the method, originally developed for ODEs, naturally extends to

$$f(\frac{1}{h_n} \sum_{i=0}^{k} \alpha_{ni} d(x_{n-i}, t_{n-i}), x_n, t_n) = 0, \tag{6.2}$$

for $n \geq k$, whereby k denotes the BDF method's order, and thus the number of previously calculated steps whose approximations are used, and α_{ni} are the BDF-coefficients, i. e.

$$\alpha_{ni} = \frac{t_n - t_{n-1}}{t_n - t_{n-i}} \prod_{j=1, j\neq i}^{k} \frac{t_n - t_{n-j}}{t_{n-i} - t_{n-j}}, \qquad \text{for } i = 1, ..., k,$$

$$\alpha_{n0} = -\sum_{i=1}^{k} \alpha_{ni}$$

Note, the order k shall not exceed 6 since otherwise the methods are not stable anymore, see e. g. [HW91]. For $k > 1$, the approximations $x_1, \ldots, x_{k-1}$ are assumed to be given with sufficient precision, for example being calculated by one-step methods of same order. Note that (6.2) itself is a nonlinear equation system in x_n, i. e.

$$f_n^{\mathrm{BDF}}(x_n) := f(\frac{1}{h_n} \sum_{i=0}^{k} \alpha_{ni} d(x_{n-i}, t_{n-i}), x_n, t_n) = 0, \tag{6.3}$$

which can be solved for instance using *Newton's method* or quasi-Newton methods such as *Broyden's method* or others, see e. g. [Hab04].

Example 6.1 (Euler's methods)
The implicit Euler method is the BDF method of order 1. For the DAE in Example 2.3 it reads for constant step size h

$$\begin{aligned} \frac{x_{n,1} - x_{n-1,1}}{h} + x_{n,1} - x_{n,2} &= 0, \\ x_{n,1} &= q_2(t_n). \end{aligned}$$

However, the explicit version of Euler's method, i. e.

$$\begin{aligned} \frac{x_{n,1} - x_{n-1,1}}{h} + x_{n-1,1} - x_{n-1,2} &= 0, \\ x_{n-1,1} &= q_2(t_{n-1}), \end{aligned}$$

does not provide an expression for $x_{n,2}$ which makes this method useless as such.

Runge-Kutta Methods *Runge-Kutta* (RK) *methods* are one-step methods which are distinguished by their coefficients, usually stored in a so-called Butcher tableau

$$\begin{array}{c|c} c & \mathcal{A} \\ \hline & b^\top \end{array} := \begin{array}{c|ccc} c_1 & a_{1,1} & \cdots & a_{1,s} \\ \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s,1} & \cdots & a_{s,s} \\ \hline & b_1 & \cdots & b_s \end{array}$$

for some stage $s \in \mathbb{N}$ larger than 1, see e. g. [But87]. If $\mathcal{A}$ is strictly lower triangular, the method is said to be explicit, implicit otherwise. There exist certain criterion for the tableau defining the method's order. Assuming the existence of the $\mathcal{A}^{-1} =: (\alpha_{ij})_{i,j=1,\dots,s}$, the Runge-Kutta methods transfer to DAEs with nonlinear derivative term as follows, see e. g. [LMT13]. For stages $i = 1, \dots, s$, we define $t_{ni} := t_{n-1} + c_i h_n$ and introduce the stage approximations X_{ni} of the solution $x(t_{ni})$. The stage approximations are computed by

$$f([DX]'_{ni}, X_{ni}, t_{ni}) = 0, \qquad \text{for } i = 1, \dots, s \tag{6.4}$$

whit derivative approximations

$$[DX]'_{ni} := \frac{1}{h_n} \sum_{j=1}^{s} \alpha_{ij} (d(X_{nj}, t_{nj}) - d(x_{n-1}, t_{n-1}))$$

yielding $x_n := X_{ns}$. Note that similar to (6.2), (6.4) is a nonlinear equation system, but the dimension scales with the method's stages, i. e.

$$f_n^{\mathrm{RK}}\left(\begin{pmatrix} X_{n1} \\ \vdots \\ X_{ns} \end{pmatrix}\right) := \begin{pmatrix} f(\frac{1}{h_n} \sum_{j=1}^{s} \alpha_{1j} (d(X_{nj}, t_{nj}) - d(x_{n-1}, t_{n-1})), X_{n1}, t_{n1}) \\ \vdots \\ f(\frac{1}{h_n} \sum_{j=1}^{s} \alpha_{sj} (d(X_{nj}, t_{nj}) - d(x_{n-1}, t_{n-1})), X_{ns}, t_{ns}) \end{pmatrix} = 0. \tag{6.5}$$

Therefore, we expect more difficulties in further solving procedures when dealing with large numbers of unknowns.

6.2.2 Newton's Method

Newton's method is an iterative method to find the root of a real or vector valued function $\bar{f} : \mathbb{R}^m \to \mathbb{R}^m$, that is an $\bar{x}$ so taht $\bar{f}(\bar{x}) = 0$, see [QSS10]. Hereby the k-th approximation $\bar{x}^{[k]}$, for $k \in \mathbb{N}$, is supposed to be a better approximation to the root than its predecessors and given by

$$\bar{x}^{[k]} = \bar{x}^{[k-1]} - J_{\bar{f}}^{-1}(\bar{x}^{[k-1]})\bar{f}(\bar{x}^{[k-1]}) \tag{6.6}$$

with initial root approximation $\bar{x}^{[0]} \in \mathbb{R}^m$ and nonsingular *Jacobian* $J_{\bar{f}}(y) = \frac{\mathrm{d}}{\mathrm{d}y}\bar{f}(y) \in \mathbb{R}^{m\times m}$.

When used for time integration, one typically chooses the previous time integration solution as an initial guess, i. e.

- $\bar{x}^{[0]} = x_{n-1}$ if $\bar{f} = f_n^{\mathrm{BDF}}$ for BDF;
- $\bar{x}^{[0]} = (X_{n-1,1}, \ldots, X_{n-1,s})$ if $\bar{f} = f_n^{\mathrm{RK}}$ for RK.

The procedure repeats till a sufficient root approximation is found, usually till $K \in \mathbb{N}$ so that

$$\left\|\bar{f}(\bar{x}^{[K]})\right\| < \epsilon \quad \text{and} \quad \left\|\bar{x}^{[K]} - \bar{x}^{[K-1]}\right\| < \delta$$

with thresholds ϵ and $\delta \in \mathbb{R}$, and for some vector norm $\|\cdot\|$. Given that the original problem is quite high dimensional, that is $m \gg 1$, one might encounter difficulties when dealing with dense matrices.

Example 6.2
Let $m = 500\,000$. A dense matrix $M \in \mathbb{R}^{m\times m}$ with the data type `double` *uses*

$$500\,000^2 \cdot 64 = 16 \times 10^{12}$$

bits of RAM, which are 2 TB.

Example 6.2 shows, that physical boundaries of computers are easily exceeded when using Newton's methods in form (6.6). The nonlinear rootfinding problem is translated into multiple solvings of linear systems

$$J_{\bar{f}}(\bar{x}^{[k-1]})z^{[k]} = -\bar{f}(\bar{x}^{[k-1]}) \tag{6.7}$$

which are equivalent to (6.6) with $z^{[k]} := \bar{x}^{[k]} - \bar{x}^{[k-1]}$.

Remark 6.3
In case (6.1) describes a linear DAE, that is of form

$$M\frac{\mathrm{d}}{\mathrm{d}t}x + Bx = q(t),$$

the functions f_n^{BDF} and f_n^{RK} are linear as well. Hence, the according root finding problems (6.3) and (6.5) are linear systems of equations which are then equal to their first step of Newton's method meaning that the Newton's method only requires one iteration.

Consider the field/circuit coupled system (5.23) reading

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA1}}(\mathbf{x}),\mathbf{x},t) = c_{\mathrm{MNA1.1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}),\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u},\mathbf{x},\mathbf{u}), \tag{6.8a}$$

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_{\mathrm{E}}(\mathbf{u}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}),\mathbf{x}). \tag{6.8b}$$

Applying BDF time integration, yields for the k-th Newton iteration in the root finding process of the n-th step (6.3) the following Jacobian:

$$\begin{aligned} J_{f_n^{\mathrm{BDF}}}(x_n^{[k]}) = \frac{\alpha_{n,0}}{h_n}&\begin{bmatrix} A_{\mathrm{C}}q_{\mathrm{C}}'(A_{\mathrm{C}}^\top\mathbf{e}_n,t_n)A_{\mathrm{C}}^\top + A_{\mathcal{E}}M_\varepsilon A_{\mathcal{E}}^\top & 0 & 0 & A_{\mathcal{E}}M_\varepsilon G & 0 & A_{\mathcal{E}}M_\varepsilon \\ 0 & \phi_{\mathrm{L}}'(\mathbf{i}_{\mathrm{L}n},t_n) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \widetilde{S}M_\varepsilon G & 0 & 0 \\ M_\varepsilon A_{\mathcal{E}}^\top & 0 & 0 & M_\varepsilon G & M_\sigma & M_\varepsilon \\ 0 & 0 & 0 & 0 & I & 0 \end{bmatrix} \\ +&\begin{bmatrix} A_{\mathrm{R}}g_{\mathrm{R}}'(A_{\mathrm{R}}^\top\mathbf{e},t)A_{\mathrm{R}}^\top + A_{\mathcal{E}}M_\sigma A_{\mathcal{E}}^\top & A_{\mathrm{L}} & A_{\mathrm{V}} & A_{\mathcal{E}}M_\sigma G & 0 & A_{\mathcal{E}}M_\sigma \\ -A_{\mathrm{L}}^\top & 0 & 0 & 0 & 0 & 0 \\ -A_{\mathrm{V}}^\top & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & H & 0 \\ M_\sigma A_{\mathcal{E}}^\top & 0 & 0 & M_\sigma G & \widetilde{C}M_\nu C & 0 \\ 0 & 0 & 0 & 0 & 0 & -I \end{bmatrix} \end{aligned} \tag{6.9}$$

whereby $x_n = (\mathbf{x}_n,\mathbf{u}_n) = (\mathbf{e}_n,\mathbf{i}_{\mathrm{L}n},\mathbf{i}_{\mathrm{V}n},\boldsymbol{\phi}_n,\mathbf{a}_n,\boldsymbol{\pi}_n)$. In case of linear characteristic functions $q_{\mathrm{C}}, g_{\mathrm{R}}$ and ϕ_{L}, the Jacobian stays unchanged for all steps n if the step size h_n is constant. Note further that due to the EM device's system properties, that is Assumption 4.20, we have symmetric positive expressions $\widetilde{S}M_\varepsilon G$ and M_ε. Contrary, $\widetilde{C}M_\nu C$ is in general to be assumed singular.

6.2.3 Linear Solving Methods

In order to solve a linear system of equations, such as (6.7), there are basically two main strategies. Either by direct methods or iterative methods, see e. g. [QSS10]. However, knowing a little bit more about the systems structure, it might be possible to combine the advantageous of multiple methods.

Let $J \in \mathbb{R}^{m\times m}$ be nonsingular and $b \in \mathbb{R}^m$. We know that the equation

$$Jz = b \tag{6.10}$$

has a unique solution $z \in \mathbb{R}^m$. In fact it is given by $z = J^{-1}b$ but, as already discussed in Example 6.2, calculating J^{-1} is avoided.

Direct Solvers Direct solvers have the advantage that their procedure ends in a predictable finite number of steps and that they are at least theoretically exact. In practice we know that they are sensitive to rounding errors, which is why they are sometimes refined by iterative schemes in the end, see e. g. [Ske80]. Further, depending on the problems' dimensions, direct solvers might encounter difficulties with limited RAM. Anyway, Gauss-Jordan elimination with pivoting is a quite common candidate for direct solvers.

Iterative Solvers Iterative solvers, on the other hand, such as *generalized minimal residual method* (GMRES) or other Krylov subspace methods, use matrix-matrix or matrix-vector multiplications only and operate sparse in RAM. Additionally, their current status can be tracked at any time and their gradually approach to the solution can be interrupted if a certain accuracy threshold is met. However, their convergence is generally not guaranteed and they are very sensitive to ill-conditioned problems. Usually, they come along with preconditioning methods. For the numerical benchmarks in Chapter 7 we will use GMRES with and without incomplete LU preconditioner as they were considered for coupled electromagnetic simulation in [Che+11] and are suitable also for non-symmetric systems.

Hybrid Solvers For instance to combine the advantages from both worlds, one can think of *hybrid solving methods* which solve a part of the linear system with one method and the other part of the system with another method. This approach might be especially useful when considering coupled systems where there are well suited linear solvers for the respectively part. It performs similar as the Schur complement reduction in [BGL05].

In particular, let the linear system of equations (6.10) be of block matrix form

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \\ s \end{pmatrix} \tag{6.11}$$

whereby A and D are assumed to be nonsingular. By the *Schur complement*, we know that a solution to (6.11) satisfies

$$x = (A - BD^{-1}C)^{-1}(r - BD^{-1}s), \tag{6.12}$$

$$y = D^{-1}(s - Cx). \tag{6.13}$$

Now, consider that we have specialized solvers to calculate linear equation systems involving A and D, respectively. With the latter solvers, we calculate solutions z^* and Y^* so that

$$Dz^* = s, \tag{6.14}$$

$$DY^* = C \tag{6.15}$$

whereby the latter is understood to be computed column wise. By the first solver we calculate the solution x^* satisfying

$$(A - BY^*)x^* = r - Bz^* \tag{6.16}$$

and again by the second solver y^* so that.

$$Dy^* = s - Cx^*. \tag{6.17}$$

It holds that (6.14) and (6.15) are equivalent to $z^* = D^{-1}s$ and $Y^* = D^{-1}C$ which, after inserting $x^* = (A - BY^*)^{-1}(r - Bz^*)$ in (6.16), yields

$$x^* = (A - BD^{-1}C)^{-1}(r - BD^{-1}s).$$

Moreover, (6.17) is equivalent to

$$y^* = D^{-1}(s - Cx^*).$$

In other words x^* and y^* fulfill (6.12) and are the solution to (6.11).

Remark 6.4
Let C be a sparse coupling block with $m \in \mathbb{N}$ non zero rows. Then there are only $m + 2$ linear equation systems involving D to solve, that is one time in (6.14), m times in (6.15) and a last time in (6.17).

In case the D matrix is the one corresponding to the refined models subsystem, e. g. the left-hand sides of the field/circuit coupled systems (5.21a), (5.22a), (5.23a), (5.24a), (5.25a), (5.26a), (5.27a) or (5.28a), then a specialized solver for the refined model equations can be used. Note that due to the circuit structure, the nonzero columns in the coupling matrix C are bounded as of the following Lemma and Corollary.

Lemma 6.5
Consider the field/circuit coupled problem (6.8) fulfilling Assumptions 3.7, 3.8, 4.20, 5.1 and 5.2. Let the Jacobian (6.9), during the time integration using BDF and Newton's method, be split up as

$$J_{f_n^{\mathrm{BDF}}}(x_n^{[k]}) = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

with A, B, C and D being each a 3×3 block. Then, C contains maximal as many nonzero columns as there are branches induced by the EM devices, i. e. m_{E}.

Proof. According to Lemma 5.6 the EM devices fit the mock element description (3.5) and (3.6). Therefore, the partial Jacobian C has only dependencies on the node potentials $\mathbf{e}$ that are not in the kernel of A_{E}. Since the ground node is the reference node among each EM device, as of Assumption 5.1, each EM device has at most as many different adjacent nodes as it has branches. Since multiple EM devices can share a node potentials, the number of nonzero columns can be smaller then m_{E}. □

Corollary 6.6
The statement of Lemma 6.5 also applies for the coupled field/circuit systems (5.21), (5.22), (5.24), (5.25), (5.26), (5.27) and (5.28) given the same Assumptions and if, for respectively (5.26), (5.27) and (5.28), the constitutive equations (5.26b), (5.27b) and (5.28b) are assigned to A and B.

Proof. Since the refined model's subsystems (5.21a), (5.22a), (5.23a), (5.24a), (5.26a), (5.27a) and (5.28a) are identical to (5.25a), their corresponding parts of the Jacobian, that is C and D, do not change. In case for (5.26), (5.27) and (5.28), C gets extra columns for $\mathbf{i}_\mathrm{E}$ which are empty. □

Further, EM devices that do not share any circuit node with another EM device, can be solved independently which enables a high potential for parallel solving. To be more specific, if

$$C = \begin{bmatrix} C_1 & & \\ & \ddots & \\ & & C_l \end{bmatrix}, \quad \text{and} \quad D = \begin{bmatrix} D_1 & & \\ & \ddots & \\ & & D_l \end{bmatrix}$$

are block diagonal, then respectively (6.14),(6.15) and (6.17) result each in l independent linear systems of equations

$$\begin{aligned} D_i z_i^* &= s_i, && \text{for } i = 1, \dots, l\ , \\ D_i Y_i^* &= C_i, && \text{for } i = 1, \dots, l\ , \\ D_i y_i^* &= s_i - C_i x_i^*, && \text{for } i = 1, \dots, l \end{aligned}$$

where $s = (s_1, \dots, s_l)$ and the overall solutions read $z^* = (z_1^*, \dots, z_l^*)$, $Y^* = (Y_1^*, \dots, Y_l^*)$ and $y^* = (y_1^*, \dots, y_l^*)$.

6.2.4 Scaling Issues

Especially when dealing with physical models, scaling becomes an important tool for numerical numerical simulation of DAEs and ODEs. There are various purposes of this technique, see [LP16]:

(i) Make independent and dependent variables dimensionless.

(ii) Make the size independent and dependent variables about unity.

(iii) Reduce the number of dependent physical parameters in the model.

With focus on the former two points, the basic idea of scaling is to introduce, for any variable q, a dimensionless variable $\bar{q}$ which takes over the role. Usually they comply the relation

$$\bar{q} = \frac{q - q_0}{q_s}$$

whereby q_0 is a reference value of q and q_s the scale of $|q|$. A common choice of the reference value is zero. Depending on what we apply this approach, we speak of *time scaling*, *equation scaling*, *variable scaling* etc. With the help of these techniques, the resulting systems are expected to be more harmless to numerical solving.

Variable Scaling Focusing on the latter technique, consider the linear DAE with time invariant matrices and sufficiently smooth right-hand side:

$$M\frac{\mathrm{d}}{\mathrm{d}t}x + Bx = q(t).$$

When applying numerical methods, such as the implicit Euler (first order BDF method) with sufficient small constant time step size h, we end up solving the linear equation

$$(\frac{1}{h}M + B)x_{n+1} = q(t_{n+1}) + Mx_n$$

where $x_n \approx x(t_n)$ is the discrete approximation of x at time point t_n. Assuming the matrix $J := (\frac{1}{h}M + B)$ to be nonsingular, we can calculate its condition $\kappa(J)$. Some applications, such as electromagnetic simulation, give rise to an extremely high condition number since the matrix entries can vary between 10^{-27} and 10^{2} usually [Mat+19]. One way to address this problem is to solve the DAE and thus the resulting linear equation system for a dimensional less version of x, namely $\bar{x}$. Neglecting the reference values, we define $\bar{x}$ via $x = S_x\bar{x}$ with a diagonal inverse scaling matrix S_x yielding

$$M(S_x\bar{x})' + BS_x\bar{x} = q(t).$$

Then, for each time step, one solves

$$JS_x\bar{x}_{n+1} = q(t_{n+1}) + MS_x\bar{x}_i$$

where S_x had been chosen such that $\kappa(JS_x) \ll \kappa(J)$. At the end we recover x_n by left-multiplying the auxiliary variable $\bar{x}_n$ with S_x.

The above theory can be extended to DAEs with nonlinear derivative terms (2.2). Here, the variable scaling works similar meaning we have to solve the DAE for auxiliary functions $\bar{f}$ and $\bar{d}$ of $\bar{x}$:

$$\bar{f}(\frac{\mathrm{d}}{\mathrm{d}t}\bar{d}(\bar{x},t),\bar{x},t) = 0$$

with

$$\begin{aligned}\bar{f}(y,\bar{x},t) &:= f(y,S_x\bar{x},t),\\ \bar{d}(\bar{x},t) &:= d(S_x\bar{x},t).\end{aligned}$$

where $x = S_x\bar{x}$. During time integration we then make use of the partial derivatives:

$$\begin{aligned}\frac{\partial}{\partial\bar{x}}\bar{d}(\bar{x},t) &= \frac{\partial}{\partial\bar{x}}d(S_x\bar{x},t) = \frac{\partial}{\partial x}d(x,t)\cdot S_x,\\ \frac{\partial}{\partial\bar{x}}\bar{f}(y,\bar{x},t) &= \frac{\partial}{\partial\bar{x}}f(y,S_x\bar{x},t) = \frac{\partial}{\partial x}f(y,x,t)\cdot S_x,\\ \frac{\partial}{\partial y}\bar{f}(y,\bar{x},t) &= \frac{\partial}{\partial y}f(y,S_x\bar{x},t) = \frac{\partial}{\partial y}f(y,x,t).\end{aligned}$$

Finally x is recovered by left-multiplying $\bar{x}$ with S_x. Note that an initial value has to be scaled accordingly. For the here considered coupled field/circuit problems we use the scales given in Table 6.1 as they are also used by `DevEM`.

quantity	variable	scale
node potential	$\mathbf{e}$	0.025852151443658
branch voltage	$\mathbf{v}$	0.025852151443658
branch current	$\mathbf{i}$	$1.91504142191233 \times 10^{-8}$
electric scalar potential	$\boldsymbol{\phi}$	0.025852151443658
magnetic vector potential	$\mathbf{a}$	$3.09003612953325 \times 10^{-07}$
quasi canonical momentum	$\boldsymbol{\pi}$	2162.86705478354

Table 6.1: Quantity scales.

Time Scaling The time scaling technique varies a little bit, but we start with the same strategy. Let $s_t \in \mathbb{R}$ be a scale. We then define τ such that $t = s_t\tau$. For simplicity we define $\tilde{x}(\tau) := x(s_t\tau) = x(t)$. It remains to solve the auxiliary system

$$\tilde{f}(\frac{\mathrm{d}}{\mathrm{d}\tau}\tilde{d}(\tilde{x}(\tau),\tau),\tilde{x}(\tau),\tau) = f(\frac{\mathrm{d}}{\mathrm{d}t}d(x(t),t),x(t),t) = 0$$

with the new functions

$$\begin{aligned}\tilde{d}(\bar{x},\tau) &:= d(\bar{x},s_t\tau),\\ \tilde{f}(\bar{y},\bar{x},\tau) &:= f(\frac{1}{s_t}\bar{y},\bar{x},s_t\tau).\end{aligned}$$

Row Scaling With row scaling we refer to a special case of equation scaling where the underlying liner system is scaled. Consider the nonsingular quadratic linear equation system

$$Jz = b$$

as it arises for instance from Newton's method (6.7) after applying time integration methods. Especially when solved with iterative methods, it is advantageous if instead a scaled system

$$S_rJz = S_rb$$

is solved whereby S_r is a diagonal nonsingular matrix. For the here considered coupled problems, we use the row scaling matrix defined by

$$S_r = \mathrm{diag}(s_i)_{i=1,\dots,m}, \quad s_i := \frac{1}{\max_{j=1,\dots,m}\{|J_{i,j}|\}}.$$

These techniques may have a positive impact on the convergence speed of the iterative solving process of the underlying linear systems of equations, see for instance the benchmark in Table 7.2 of Example 7.1.

6.2.5 Structural Exploitation

If one has a priori structural knowledge of the system to be solved, one can use this knowledge to construct tailored numerical methods to solve it. In the following we present strategies on how to improve the time integration process of the field/circuit coupled systems, following the approach in [Vaa96; Sch11].

During time integration, we assume to encounter linear systems of equations (6.7) which are of the form

$$\begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & D & J_{23} \\ J_{31} & J_{32} & J_{33} \end{bmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \tag{6.18}$$

where D is a diagonal nonsingular matrix. Note that D^{-1} can be calculated in linear time which enables the opportunity to transform (6.18) in a similar fashion to the Schur complement just before the linear solver is applied. First, we figure that

$$z_2 = D^{-1}(b_2 - J_{21}z_1 - J_{23}z_3) \tag{6.19}$$

which after inserting in (6.18) yields

$$\begin{bmatrix} J_{11} - J_{12}D^{-1}J_{21} & J_{13} - J_{12}D^{-1}J_{23} \\ J_{31} - J_{32}D^{-1}J_{21} & J_{33} - J_{32}D^{-1}J_{23} \end{bmatrix} \begin{pmatrix} z_1 \\ z_3 \end{pmatrix} = \begin{pmatrix} b_1 - J_{12}D^{-1}b_2 \\ b_3 - J_{32}D^{-1}b_2 \end{pmatrix}. \tag{6.20}$$

Next, we apply the linear solver to (6.20), yielding z_1 and z_3, followed by a recovering process of z_2 using (6.19).

Eliminating the Quasi-Canonical Momentum In order to avoid second order time derivative of the Lorenz-gauged $\mathbf{A} - \boldsymbol{\varphi}$ formulation (4.18), as a model for the electromagnetic device problem, we make use of the quasi-canonical momentum (4.19), i. e.

$$\boldsymbol{\Pi} = \frac{\partial}{\partial t}\mathbf{A}$$

as an auxiliary variable alongside the magnetic vector potential $\mathbf{A}$ and electric scalar potential $\boldsymbol{\varphi}$. After spatial discretization in Section 4.5, these quantities are translated to the unknown vector functions $\mathbf{a}$ and $\boldsymbol{\pi}$, associated with links, and $\boldsymbol{\phi}$, associated with mesh points on the computational mesh. The degrees of freedom for $\mathbf{a}$ and $\boldsymbol{\pi}$ match each other and are a multiple of the ones for $\boldsymbol{\phi}$. Given a Cartesian mesh, for instance, we can assume approximatively three times more links then mesh points.
Now, we can exploit the structural information of the quasi-canonical momentum in that we eliminate either $\mathbf{a}$ or $\boldsymbol{\pi}$ as variables in a lower part of the solving hierarchy. To be more precise, recall for instance the previously considered field/circuit coupled system (6.8). The system's Jacobian encountered during BDF time integration with Newton's method are given by (6.9).

In the block row of the EM subsystem (5.8) we encounter for the discrete quasi-canonical momentum two diagonal matrices yielding the relation

$$\mathbf{a} = \frac{h_n}{\alpha_{n,0}}\boldsymbol{\pi}. \tag{6.21}$$

Inserting $\mathbf{a}$ into the remaining equations, as in (6.20), yields the new linear system matrix reduced by $\mathbf{a}$

$$\frac{\alpha_{n,0}}{h_n}\begin{bmatrix} A_{\mathrm{C}}q'_{\mathrm{C}}(A_{\mathrm{C}}^\top \mathbf{e}_n, t_n)A_{\mathrm{C}}^\top + A_{\mathcal{E}}M_\varepsilon A_{\mathcal{E}}^\top & 0 & 0 & A_{\mathcal{E}}M_\varepsilon G & A_{\mathcal{E}}M_\varepsilon \\ 0 & \phi'_{\mathrm{L}}(\mathbf{i}_{\mathrm{L}n}, t_n) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \widetilde{S}M_\varepsilon G & 0 \\ M_\varepsilon A_{\mathcal{E}}^\top & 0 & 0 & M_\varepsilon G & M_\varepsilon \end{bmatrix}$$
$$+\begin{bmatrix} A_{\mathrm{R}}g'_{\mathrm{R}}(A_{\mathrm{R}}^\top \mathbf{e}, t)A_{\mathrm{R}}^\top + A_{\mathcal{E}}M_\sigma A_{\mathcal{E}}^\top & A_{\mathrm{L}} & A_{\mathrm{V}} & A_{\mathcal{E}}M_\sigma G & A_{\mathcal{E}}M_\sigma \\ -A_{\mathrm{L}}^\top & 0 & 0 & 0 & 0 \\ -A_{\mathrm{V}}^\top & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{h_n}{\alpha_{n,0}}H \\ M_\sigma A_{\mathcal{E}}^\top & 0 & 0 & M_\sigma G & \frac{h_n}{\alpha_{n,0}}\widetilde{C}M_\nu C + M_\sigma \end{bmatrix}.$$

The right-hand side remains unchanged but reduced by the latter block entry since the corresponding b_2 is zero. After solving the linear system, $\mathbf{a}$ is recovered by (6.21). The dimension of (6.20) is approximately $\frac{4}{7}$ times the one of (6.18), given a Cartesian mesh with a relatively small circuit subsystem. Note that for sufficient small h_n the block entry $\frac{\alpha_{n,0}}{h_n}M_\varepsilon + \frac{h_n}{\alpha_{n,0}}\widetilde{C}M_\nu C + M_\sigma$ becomes positive definite and is symmetric. If this part is solved via a hybrid solving approach, one can apply even faster iterative solvers such as *conjugate gradient* method, which is guaranteed to converge after a finite number of steps.

Remark 6.7
The transformation (6.20) *can also be applied multiple times, for example if there is more than one system or higher order Runge-Kutta methods are used. Further, it can be applied to all couped system variants* (5.21), (5.22), (5.23), (5.24), (5.25), (5.26), (5.27) *and* (5.28), *since they shared the same EM subsystem* (5.8) *as of Remark 5.11.*

All these techniques are derived using the formulations introduced in Chapter 5. With minimal structural knowledge, such as equation and variable order, they can be adopted to refined model equations obtained by the black-box modeler `DevEM`, as done in Chapter 7. An improvement of the computational time can be observed in Figure 7.3.

6.3 Waveform Relaxation Method

As we understand from application, there is a need of being capable to solve bigger problems. But where the arising dynamical systems become larger, they become more challenging. Instead of solving such a system at once, one might think of splitting it up into multiple subsystems and

solving them iteratively. In particular, solving such a subsystem means integrating numerically while using previously calculated solutions of the other subsystems as inputs. This approach is referred to as *waveform relaxation method*. Essentially, this method uses iterative relaxation schemes such as of Gauss-Seidel or Jacobi type whereby the elements of relaxation are waveforms (solutions) of the unknown variables, see [Lel82; Whi+85].
Preferably, the subsystems are chosen in such a way that they are substantially easier to solve, for example due to nice structure or by applying individually tailored numerical methods to the subsystems, cf. [AG01]. Whenever coupling different physical phenomena, their individual models might have undergone plenty of numerical analysis and may naturally yield a good starting point of subsystem choices. Indeed, for the here considered circuit coupled EM device problem, we choose the subsystems to match the individual phenomenas' modeling, as already given in (5.21)-(5.28). Splitting the overall problem as such yields an ODE for one of the EM devices subsystems and, for example, enables the usage of explicit time integration schemes. This modular approach can be of advantage when considering parallel computation but the bottleneck here is the subsystems synchronization.

Whereas an early usage for WR scheme was for integrated circuits[1] see e. g. [LRS82; WS12], the application of WR methods is by far not restricted to those. For instance, when dealing with coupled multi-physical problems, see e. g. [Cle+12]. Especially for field/circuit coupled problems, WR schemes became quite popular, see for instance [SDB10; SDB12; Bar+13].
In the early days, investigation of WR methods was also accompanied by an increased interest in accelerating the convergence thereof. In the 1960s, one used successive overrelaxation to speed up the convergence and later introduced vector extrapolation methods, see e. g. [JS95]. A new and promising approach with regard to field/circuit coupled problems is given in e. g. [Cor+17; Cor20] where a connection to optimized Schwarz WR methods is made which were earlier transferred to circuit simulation by Martin Gander and Albert Ruehli, cf. [GR04; GR10; CG13; AGR14].

If the subsystems are ODEs, then these methods are understood to converge and, provided a good initial guess and a weak coupling of ODE subsystems, one may expect the computing time to reduce significantly, see [MN87; Bur95]. In case that DAEs are involved these methods are not guaranteed to converge, see e. g. [AG01; Bar+13], if an additional contraction condition is violated. For a further insight in the coupling structure see for instance [Mie89]. Analysis concerning DAEs with index 1 are to be found for example in [Ebe04; BBS14; JK96] and for index-2 see e. g. [SZF06].

For the here considered coupled field/circuit systems of Chapter 5 we require analysis of WR schemes for DAE-ODE coupled system similar to [PT18] but for a more complex coupling scheme, that is one which allows for additional derivatives of the quantities. One special case of these coupling schemes is studied in [ST20] and, therefore, the following analysis can be considered as a continuation. We provide a convergence theorem, extending the ones in [PT18; JK96; Sau19], that applies to field/circuit coupled problems with full-wave MEs modeling of EM devices, also called *full-Maxwell*. Additionally, we provide sufficient topological criteria for convergence of both Gauss-Seidel and Jacobi type WR schemes for the full-Maxwell field/circuit coupled systems.

[1] for example in the circuit simulator `MOTIS` [CGK75]

Types of Waveform Relaxation Schemes The here presented approach is a straight forward generalization of the one for implicit nonlinear DAEs in standard form, in [Lel82], to DAEs with nonlinear derivative term. As the WR method is represented by an iterative scheme, we decorate the unknowns with $(\cdot)^{[k]}$ representing the solution after k iterations whereby $k \in \mathbb{N}$ is the iteration counter. Moreover, from hereon we omit the time argument t of the unknown vector functions for the sake of readability.

For an integer $r \geq 2$, consider a coupled system (2.6), that is

$$f(\frac{\mathrm{d}}{\mathrm{d}t}d(x,t),x,t) = \begin{pmatrix} f_1(\frac{\mathrm{d}}{\mathrm{d}t}d_1(x,t),x,t) \\ \vdots \\ f_r(\frac{\mathrm{d}}{\mathrm{d}t}d_r(x,t),x,t) \end{pmatrix} = 0, \tag{6.22}$$

together with a consistent initial condition $x(t_0) = x_0$. Motivated by [Lel82; Bur95; JK96], the general form of representing a waveform relaxation scheme for (6.22) with initial condition reads

$$F(\frac{\mathrm{d}}{\mathrm{d}t}D(x^{[k]},x^{[k-1]},t),x^{[k]},x^{[k-1]},t) = 0, \quad x^{[k]}(t_0) = x_0 \tag{6.23}$$

with an initial guess $x^{[0]}$, also complying $x^{[0]}(t_0) = x_0$, where $F : \mathbb{R}^{n_1+\cdots+n_r} \times \mathcal{D} \times \mathcal{D} \times \mathcal{I} \to \mathbb{R}^{m_1+\cdots+m_r}$ and $D : \mathcal{D} \times \mathcal{D} \times \mathcal{I} \to \mathbb{R}^{n_1+\cdots+n_r}$ satisfying

$$\begin{aligned} F(\frac{\mathrm{d}}{\mathrm{d}t}D(x,x,t),x,x,t) &= (f_i(\frac{\mathrm{d}}{\mathrm{d}t}d_i(x,t),x,t))_{i=1,\ldots,r} \\ D(x,x,t) &= (d_i(x,t))_{i=1,\ldots,r}. \end{aligned}$$

In this treatise, the focus will be upon two popular WR schemes. The first one is of Gauss-Seidel type where the subsystems are solved sequentially one by another. The second WR scheme is of Jacobi type where the subsystems are solved alike independent of the order. Whereas Gauss-Seidel type schemes have usually better convergence properties then Jacobi type, the latter one can be solved in parallel.
Note that these WR types are by far not the only ones. Alternatives are, to mention a few, the Picard iteration or the successive over-relaxation method, see for instance [MN87; Bur95; JK96].

Given a coupled system with consistent initial condition, such as (6.22), we now introduce the Gauss-Seidel type waveform relaxation scheme by the following definition.

Definition 6.8 (Gauss-Seidel type WR scheme)
The waveform relaxation scheme (6.23) *is said to be of Gauss-Seidel type if, respectively,*

$$\frac{\partial}{\partial v}F(u,v,w,t)|_{u=\frac{\mathrm{d}}{\mathrm{d}t}D(v,w,t),v=x,w=x} \quad \textit{and} \quad \frac{\partial}{\partial v}D(v,w,t)|_{v=x,w=x}$$

are lower block triangular matrices consisting of the lower block triangular part of

$$\frac{\partial}{\partial x}(f_i(\frac{\mathrm{d}}{\mathrm{d}t}d_i(x,t),x,t))_{i=1,\ldots,r}. \quad \textit{and} \quad \frac{\partial}{\partial x}(d_i(x,t))_{i=1,\ldots,r}.$$

The second class of consideration in this treatise, the Jacobi type waveform relaxation scheme, is introduced similarly.

Definition 6.9 (Jacobi type WR scheme)
The waveform relaxation scheme (6.23) *is said to be of Jacobi type if*

$$\frac{\partial}{\partial v} F(u,v,w,t)|_{u=\frac{\mathrm{d}}{\mathrm{d}t}D(v,w),v=x,w=x} = \operatorname{blockdiag}\left(\frac{\partial}{\partial x}(f_i(\frac{\mathrm{d}}{\mathrm{d}t}d_i(x,t),x,t))_{i=1,\dots,r}\right)$$

and

$$\frac{\partial}{\partial v} D(v,w,t)|_{v=x,w=x} = \operatorname{blockdiag}\left(\frac{\partial}{\partial x}(d_i(x,t))_{i=1,\dots,r}\right).$$

Windowing Technique In the following we introduce briefly the so-called *windowing technique*. Consider the coupled system (6.22) complying the initial condition $x(t_0) = x_0$. Instead of solving the subsystems on the whole time interval $[t_0, T] \subset \mathbb{R}$ during each sweep of the WR scheme, one can also think of solving them separately on windows $[T_n, T_{n+1}]$ where the T_n are the so-called synchronization points complying

$$t_0 = T_0 < T_1 < \cdots < T_N = T,$$

for $N \in \mathbb{N}$ windows, see e. g. [AG01]. We start by applying a WR scheme for $\mathcal{I} = [T_0, T_1]$ and then continue with the next window till $\mathcal{I} = [T_{N-1}, T_N]$ was processed whereby the last sweeps' approximation at the end of each window translates to the initial condition for the next one. This technique was also used for instance in [Bar+13] or [SDB10] with additional sweep control showing a significant efficiency increase in convergence on the whole time interval.

6.3.1 Gauss-Seidel Type for Field/Circuit Coupled System

We introduce Gauss-Seidel type waveform-relaxation schemes for the incorporated, shifted and black-box versions of the field/circuit coupled systems in Section 5. Observe that in accordance with Definition 6.8, the Gauss-Seidel WR scheme for two subsystems reads

$$f_1(\frac{\mathrm{d}}{\mathrm{d}t}d_1(u^{[k]},t),u^{[k]},t) = c_1(\frac{\mathrm{d}}{\mathrm{d}t}d_{1-2}(x^{[k-1]},t),x^{[k-1]},t), \tag{6.24a}$$

$$f_2(\frac{\mathrm{d}}{\mathrm{d}t}d_2(x^{[k]},t),x^{[k]},t) = c_2(\frac{\mathrm{d}}{\mathrm{d}t}d_{1-2}(x^{[k]},t),\frac{\mathrm{d}}{\mathrm{d}t}u^{[k]},x^{[k]},u^{[k]},t) \tag{6.24b}$$

from top to bottom with $k \in \mathbb{N}$ being the iteration counter.

Let $\mathbf{x}^{[0]}$ and, consequently, $\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[0]})$ be initial guesses on the time interval $\mathcal{I} = [t_0, T] \subset \mathbb{R}$, so that $\mathbf{x}^{[0]}(t_0) = x_0$. All equations are completed with their accordingly consistent initial values, i. e. $\mathbf{u}^{[k]}(t_0) = u_0$, $\mathbf{x}^{[k]}(t_0) = x_0$ and, if necessary, $\mathbf{i}_{\mathrm{E}}^{[k]}(t_0) = i_{\mathrm{E}0}$.

Incorporated Versions Deduced from (6.24), the Gauss-Seidel type WR schemes of (5.21), (5.22) and (5.23), while using the variable splitting $\mathbf{u}$ and $\mathbf{x}$ as well as the same order, read

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \tag{6.25a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA1.0}}(\mathbf{u}^{[k]}), \tag{6.25b}$$

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \tag{6.26a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA1.1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{x}^{[k]}, \mathbf{u}^{[k]}), \tag{6.26b}$$

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \tag{6.27a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA1.2}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{x}^{[k]}, \mathbf{u}^{[k]}). \tag{6.27b}$$

Lemma 6.10
Given the common initial guess $\mathbf{x}^{[0]}$, then, for any $k \in \mathbb{N}$, we obtain the following equivalence statements:

(i) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ is a solution of (6.26) if and only if it is a solution of (6.27).

Proof. Given the same initial guess $\mathbf{x}^{[0]}$, after the first iteration sweep of (6.26), that is for $k = 1$, we obtain $\mathbf{u}^{[1]}$ and $\mathbf{x}^{[1]}$ solving

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[1]} + b_{\mathrm{E}}(\mathbf{u}^{[1]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[0]}), \mathbf{x}^{[0]}),$$

which is the very same as (6.27a), and

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA1}}(\mathbf{x}^{[1]}), \mathbf{x}^{[1]}, t) = c_{\mathrm{MNA1.1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[1]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[1]}, \mathbf{x}^{[1]}, \mathbf{u}^{[1]}). \tag{6.28}$$

Since (6.28) is the same as (6.27a) it has the same solution $\mathbf{u}^{[1]}$. Especially it is $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{a}^{[1]} = \boldsymbol{\pi}^{[1]}$, see the according subsystem definition given in (5.8). Furthermore, by definition of $c_{\mathrm{MNA1.1}}$ in (5.19b) and $d_{\mathrm{E-MNA}}$ in (5.9), we have

$$\begin{aligned}
&c_{\mathrm{MNA1.1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[1]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[1]}, \mathbf{x}^{[1]}, \mathbf{u}^{[1]}) \\
&= \begin{pmatrix} -A_{\mathrm{M}}\Lambda^{\top}G_{\Gamma}^{\top}\left[M_{\varepsilon}\left(G\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}^{[1]} + \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\pi}^{[1]}\right) + \frac{\mathrm{d}}{\mathrm{d}t}(M_{\varepsilon}G_{\Gamma}\Lambda A_{\mathrm{E}}^{\top}\mathbf{e}^{[1]}) + M_{\sigma}\left(G\boldsymbol{\phi}^{[1]} + G_{\Gamma}\Lambda A_{\mathrm{E}}^{\top}\mathbf{e}^{[1]} + \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{a}^{[1]}\right)\right] \\ 0 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} -A_{\mathrm{M}}\Lambda^{\top}G_{\Gamma}^{\top}\left[M_{\varepsilon}\left(G\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}^{[1]} + \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\pi}^{[1]}\right) + \frac{\mathrm{d}}{\mathrm{d}t}(M_{\varepsilon}G_{\Gamma}\Lambda A_{\mathrm{E}}^{\top}\mathbf{e}^{[1]}) + M_{\sigma}\left(G\boldsymbol{\phi}^{[1]} + G_{\Gamma}\Lambda A_{\mathrm{E}}^{\top}\mathbf{e}^{[1]} + \boldsymbol{\pi}^{[1]}\right)\right] \\ 0 \\ 0 \end{pmatrix} \\
&= c_{\mathrm{MNA1.2}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[1]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[1]}, \mathbf{x}^{[1]}, \mathbf{u}^{[1]})
\end{aligned}$$

implying that $\mathbf{x}^{[1]}$ solves (6.26b) if and only if it solves (6.27b). From that we conclude the equivalence of (6.26) and (6.27), for $k = 1$. As this concept applies also for any $k > 1$, the lemmas's statement is proofed. □

Remark 6.11
Contrary to Propositions 5.10, the systems (6.25), (6.26) and (6.27) are not entirely equivalent anymore. Whereas, for any $k \in \mathbb{N}$, $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ is a solution of (6.26) if and only if it is a solution of (6.27), see Lemma 6.10, it may not be one of (6.25) and vice versa.

The plausibility of Remark 6.11 can be checked as follows. Since (6.25a) and (6.26a) are the same equation, $\mathbf{u}^{[1]}$ solves the first one if and only if it solves the second one. More particularly, from the second row block of (6.25a) we obtain, see again (5.8),

$$\begin{aligned}
&M_\varepsilon \left(G \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\phi}^{[1]} + \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\pi}^{[1]} \right) + M_\sigma \left(G \boldsymbol{\phi}^{[1]} + \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{a}^{[1]} \right) + \widetilde{C} M_\nu C \mathbf{a}^{[1]} \\
=&\frac{\mathrm{d}}{\mathrm{d}t} (M_\varepsilon G_\Gamma \Lambda A_\mathrm{E}^\top \mathbf{e}^{[0]}) + M_\sigma G_\Gamma \Lambda A_\mathrm{E}^\top \mathbf{e}^{[0]}.
\end{aligned}$$

Multiplication from both sides with $A_\mathrm{M} \Lambda^\top G_\Gamma^\top$ yields

$$\begin{aligned}
&A_\mathrm{M} \Lambda^\top G_\Gamma^\top \widetilde{C} M_\nu C \mathbf{a}^{[1]} \\
=&A_\mathrm{M} \Lambda^\top G_\Gamma^\top \big[M_\varepsilon \left(G \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\phi}^{[1]} + \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\pi}^{[1]} \right) + \frac{\mathrm{d}}{\mathrm{d}t} (M_\varepsilon G_\Gamma \Lambda A_\mathrm{E}^\top \mathbf{e}^{[0]}) \\
&+ M_\sigma \left(G \boldsymbol{\phi}^{[1]} + G_\Gamma \Lambda A_\mathrm{E}^\top \mathbf{e}^{[0]} + \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{a}^{[1]} \right) \big] \\
=&f_{\mathrm{MNA1.1}} (\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}} (\mathbf{x}^{[0]}), \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{u}^{[1]}, \mathbf{x}^{[0]}, \mathbf{u}^{[1]}).
\end{aligned}$$

In the second part of the sweep, $\mathbf{x}^{[1]}$ is obtained by the equation

$$f_{\mathrm{MNA1}} (\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}} (\mathbf{x}^{[1]}), \mathbf{x}^{[1]}, t) = c_{\mathrm{MNA1.0}} (\mathbf{u}^{[1]})$$

which, exploiting the latter observation, is equivalent to

$$f_{\mathrm{MNA1}} (\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}} (\mathbf{x}^{[1]}), \mathbf{x}^{[1]}, t) = c_{\mathrm{MNA1.1}} (\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}} (\mathbf{x}^{[0]}), \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{u}^{[1]}, \mathbf{x}^{[0]}, \mathbf{u}^{[1]}).$$

Hence, $\mathbf{x}^{[1]}$ cannot, in general, be a solution to

$$f_{\mathrm{MNA1}} (\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}} (\mathbf{x}^{[1]}), \mathbf{x}^{[1]}, t) = c_{\mathrm{MNA1.1}} (\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}} (\mathbf{x}^{[1]}), \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{u}^{[1]}, \mathbf{x}^{[1]}, \mathbf{u}^{[1]}),$$

the second part of the first sweep of (6.26b).

Shifted Versions Analogously, for (5.24) and (5.25) we deduce, respectively, the Gauss-Seidel type iteration schemes

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \tag{6.29a}$$

$$f_{\mathrm{MNA2}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA2}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA2.1}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}), \tag{6.29b}$$

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \tag{6.30a}$$

$$f_{\mathrm{MNA2}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA2}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA2.2}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}) \tag{6.30b}$$

Lemma 6.12
Given the common initial guess $\mathbf{x}^{[0]}$*, then, for any* $k \in \mathbb{N}$*, we obtain the following equivalence statements:*

(i) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *is a solution of* (6.29) *if and only if it is a solution of* (6.30)*.*

(ii) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *is a solution of either* (6.26)*,* (6.27)*,* (6.29) *and* (6.30) *or none of them.*

Proof. For (i) we figure, that the equations (6.29a) and (6.30a) equal. Hence, for the same $\mathbf{x}^{[k-1]}$, both equations are solved by the same $\mathbf{u}^{[k]}$. In particular, it holds $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{a}^{[k]} = \boldsymbol{\pi}^{[k]}$ as of each subsystem's definition in (5.8). Now, for the solution $\mathbf{x}^{[k]}$ of (6.29b) holds

$$\begin{aligned}
f_{\mathrm{MNA2}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA2}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) &= c_{\mathrm{MNA2.1}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}) \\
&= \begin{pmatrix} -A_{\mathrm{M}}\Lambda^{\top}G_{\Gamma}^{\top}\left[M_{\varepsilon}\left(G\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}^{[k]} + \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\pi}^{[k]}\right) + M_{\sigma}\left(G\boldsymbol{\phi}^{[k]} + \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{a}^{[k]}\right)\right] \\ 0 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} -A_{\mathrm{M}}\Lambda^{\top}G_{\Gamma}^{\top}\left[M_{\varepsilon}\left(G\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}^{[k]} + \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\pi}^{[k]}\right) + M_{\sigma}\left(G\boldsymbol{\phi}^{[k]} + \boldsymbol{\pi}^{[k]}\right)\right] \\ 0 \\ 0 \end{pmatrix} \\
&= c_{\mathrm{MNA2.2}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}).
\end{aligned}$$

and thus it solves (6.30b) and vice versa. As for the induction's start, we had chosen the same initial guess $\mathbf{x}^{[0]}$ for all iteration schemes. In order to show (ii) we only have to show that, for instance, (6.30) is equivalent to (6.27), since (6.26) and (6.27) are already equivalent as of Lemma 6.10. Again, $\mathbf{u}^{[k]}$ is the common solution to both (6.30a) and (6.27a) for the same

$\mathbf{x}^{[k-1]}$. As for the solution $\mathbf{x}^{[k]}$ of (6.27b) it holds, due to (5.17),

$$
\begin{aligned}
& f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA1.2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{x}^{[k]}, \mathbf{u}^{[k]}) \\
\Leftrightarrow \quad & f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = \begin{pmatrix} -A_{\mathrm{E}}[\frac{\mathrm{d}}{\mathrm{d}t} q_{\mathrm{E}}(A_{\mathrm{E}}^{\top}\mathbf{e}^{[k]}) + g_{\mathrm{E}}(A_{\mathrm{E}}^{\top}\mathbf{e}^{[k]}) + s_{\mathrm{E2}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]})] \\ 0 \\ 0 \end{pmatrix} \\
\Leftrightarrow \quad & f_{\mathrm{MNA2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA2}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = \begin{pmatrix} -A_{\mathrm{E}} s_{\mathrm{E2}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}) \\ 0 \\ 0 \end{pmatrix} = c_{\mathrm{MNA2.2}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}).
\end{aligned}
$$

□

Black-box Versions We observe that the black-box versions, which are (5.26), (5.27) and (5.28), consist of three equations each. Keeping in mind that there are still two physical phenomena, we persuade to define two subsystems for the WR method. Whereas the EM field and MNA equations are clearly to be separated, the current coupling equations' situation is not settled. Taking into account that the current coupling equations are provided by the black-box solver, group them with the EM field equations. According to (6.24), the Gauss-Seidel type iteration schemes read, respectively for (5.26), (5.27) and (5.28) with the splitting of $(\mathbf{u}, \mathbf{i}_{\mathrm{E}})$ and $\mathbf{x}$,

$$
\left\{ \begin{aligned} M_{\mathrm{E}}\tfrac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) &= c_{\mathrm{E}}(\tfrac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \\ \mathbf{i}_{\mathrm{E}}^{[k]} &= f_{\mathrm{E0}}(\mathbf{u}^{[k]}), \end{aligned} \right. \tag{6.31a}
$$

$$
f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA}}(\mathbf{i}_{\mathrm{E}}^{[k]}), \tag{6.31b}
$$

$$
\left\{ \begin{aligned} M_{\mathrm{E}}\tfrac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) &= c_{\mathrm{E}}(\tfrac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \\ \mathbf{i}_{\mathrm{E}}^{[k]} &= f_{\mathrm{E1}}(\tfrac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \tfrac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}), \end{aligned} \right. \tag{6.32a}
$$

$$
f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA}}(\mathbf{i}_{\mathrm{E}}^{[k]}), \tag{6.32b}
$$

$$
\left\{ \begin{aligned} M_{\mathrm{E}}\tfrac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) &= c_{\mathrm{E}}(\tfrac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \\ \mathbf{i}_{\mathrm{E}}^{[k]} &= f_{\mathrm{E2}}(\tfrac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \tfrac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}), \end{aligned} \right. \tag{6.33a}
$$

$$
f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA}}(\mathbf{i}_{\mathrm{E}}^{[k]}), \tag{6.33b}
$$

Remark 6.13
Instead of solving the first two equations of each black-box variant sequentially, these subsystems can be considered as one each, that is (6.31a), (6.32a) and (6.33a). Doing so, the schemes (6.31), (6.32) and (6.33) represent as well exactly the Gauss-Seidel type waveform relaxation scheme for two subsystems with the splitting $(\mathbf{u}, \mathbf{i}_{\mathrm{E}})$ and $\mathbf{x}$.

The reason for keeping the first two equations grouped is due to the nature of black-box solvers which calculate, for given $(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{[k-1]}, \mathbf{x}^{[k-1]})$, MEs and the current coupling equation in one sweep while yielding $\mathbf{i}_{\mathrm{E}}^{[k]}$ only.

Lemma 6.14
Given the common initial guess $\mathbf{x}^{[0]}$*, all black-box iteration schemes, that is* (6.31)*,* (6.32) *and* (6.33)*, are analytically equivalent to* (6.25) *in the following manner. For any* $k \in \mathbb{N}$

(i) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]})$ *is a solution of* (6.31) *if and only if* $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *is a solution of* (6.25) *and*

$$\mathbf{i}_{\mathrm{E}}^{[k]} = f_{\mathrm{E0}}(\mathbf{u}^{[k]});$$

(ii) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]})$ *is a solution of* (6.32) *if and only if* $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *is a solution of* (6.25) *and*

$$\mathbf{i}_{\mathrm{E}}^{[k]} = f_{\mathrm{E1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{x}^{[k-1]});$$

(iii) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]})$ *is a solution of* (6.33) *if and only if* $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *is a solution of* (6.25) *and*

$$\mathbf{i}_{\mathrm{E}}^{[k]} = f_{\mathrm{E2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{x}^{[k-1]}).$$

Proof. Looking back at the proof of Lemma 6.10 and considering the induction step $k-1$ to k, we figure that after solving the first subsystem of all schemes, that is (6.31), (6.32), (6.33) and (6.25), holds

$$\begin{aligned}
&A_{\mathrm{M}}\Lambda^{\top}G_{\Gamma}^{\top}\widetilde{C}M_{\nu}C\mathbf{a}^{[k]} \\
=&A_{\mathrm{M}}\Lambda^{\top}G_{\Gamma}^{\top}\left[M_{\varepsilon}\left(G\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}^{[k]} + G_{\Gamma}\Lambda A_{\mathrm{E}}^{\top}\mathbf{e}^{[k-1]} + \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\pi}^{[k]}\right) + M_{\sigma}\left(G\boldsymbol{\phi}^{[k]} + G_{\Gamma}\Lambda A_{\mathrm{E}}^{\top}\mathbf{e}^{[k-1]} + \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{a}^{[k]}\right)\right].
\end{aligned}$$

Exploiting further that $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{a}^{[k]} = \boldsymbol{\pi}^{[k]}$, we obtain

$$\begin{aligned}
\mathbf{i}_{\mathrm{E}}^{[k]} &= f_{\mathrm{E0}}(\mathbf{u}^{[k]}) \\
&= f_{\mathrm{E1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{x}^{[k-1]}) \\
&= f_{\mathrm{E2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{x}^{[k-1]}).
\end{aligned}$$

The rest follows immediately by definition. □

For the previously introduced iteration schemes, one could change the subsystems' order. We only considered the case, where the ODE subsystem is solved first; in case of the black-box versions it is the ODE from MEs and then the current coupling equation as fixed by the black-box behavior. We expect more difficulties when solving the DAE subsystem first since the algebraic constraints do not keep the same during the solving process in this case.

Corollary 6.15
Collecting all the information from Lemmata 6.10, 6.12 and 6.14, we figure that there are two equivalence classes of Gauss-Seidel type waveform relaxation scheme applied to the field/circuit coupled systems.

GS1 *The first class contains the incorporated scheme (6.25) and all black-box variants (6.31), (6.32) and (6.33).*

GS2 *The second class contains the incorporated schemes (6.26) and (6.27) as well as all shifted variants (6.29) and (6.30).*

According to Corollary 6.15, the solutions obtained after each iteration sweep do not differ, analytically, among those within the same equivalence class, that is GS1 and GS2. In other words, each iteration schemes from (6.25), (6.31), (6.32) and (6.33) converges if and only if all of them converge and (6.26), (6.27), (6.29) and (6.30) converge if and only if all of them converge. In general, the solutions across these classes do not coincide. This behavior is inherited from Remark 6.11. Anyway, for the forthcoming convergence analysis it is sufficient to choose one representative among each class. The representatives are (6.25) for GS1 and (6.30) for GS2. Nevertheless, numerical solutions of these sequences are expected to differ at least by magnitude of machine precision.

6.3.2 Jacobi Type for Field/Circuit Coupled System

In the same fashion to the Gauss-Seidel type WR iteration schemes, we introduce the Jacobi types for the incorporated, shifted and black-box variants of field/circuit coupled systems. By Definition 6.9, the Jacobi type WR scheme for two subsystems reads

$$f_1(\frac{\mathrm{d}}{\mathrm{d}t}d_1(u^{[k]},t),u^{[k]},t) = c_1(\frac{\mathrm{d}}{\mathrm{d}t}d_{1-2}(x^{[k-1]},t),x^{[k-1]},t), \tag{6.34a}$$

$$f_2(\frac{\mathrm{d}}{\mathrm{d}t}d_2(x^{[k]},t),x^{[k]},t) = c_2(\frac{\mathrm{d}}{\mathrm{d}t}d_{1-2}(x^{[k]},t),\frac{\mathrm{d}}{\mathrm{d}t}u^{[k-1]},x^{[k]},u^{[k-1]},t) \tag{6.34b}$$

from top to bottom with $k \in \mathbb{N}$ being the iteration counter. Let $\mathbf{x}^{[0]}$, $\mathbf{u}^{[0]}$ and, consequently, $\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[0]})$ and $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[0]}$ as initial guesses on the time interval $\mathcal{I} = [t_0, T] \subset \mathbb{R}$, that also comply with the consistent initial conditions $\mathbf{x}^{[0]}(t_0) = x_0$ and $\mathbf{u}^{[0]}(t_0) = u_0$. Again, all equations are completed with their according consistent initial values, i. e. $\mathbf{u}^{[k]}(t_0) = u_0$ and $\mathbf{x}^{[k]}(t_0) = x_0$.

Incorporated Versions Deduced from (6.34), the Jacobi type WR scheme applied on (5.21), (5.22) and (5.23) reads respectively

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}),\mathbf{x}^{[k-1]}), \tag{6.35a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}),\mathbf{x}^{[k]},t) = c_{\mathrm{MNA1.0}}(\mathbf{u}^{[k-1]}), \tag{6.35b}$$

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \tag{6.36a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA1.1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}, \mathbf{x}^{[k]}, \mathbf{u}^{[k-1]}), \tag{6.36b}$$

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \tag{6.37a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA1.2}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}, \mathbf{x}^{[k]}, \mathbf{u}^{[k-1]}). \tag{6.37b}$$

Shifted Versions Analogously, for (5.24) and (5.25) we deduce, respectively, the iteration schemes

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \tag{6.38a}$$

$$f_{\mathrm{MNA2}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA2}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA2.1}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}, \mathbf{u}^{[k-1]}), \tag{6.38b}$$

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \tag{6.39a}$$

$$f_{\mathrm{MNA2}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA2}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA2.2}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}, \mathbf{u}^{[k-1]}) \tag{6.39b}$$

Black-Box Versions Keeping in mind that, due to the black-box behavior, the variables $(\mathbf{u}, \mathbf{i}_{\mathrm{E}})$ are solved simultaneously, the Jacobi type iteration scheme reads, for respectively (5.26), (5.27) and (5.28),

$$\begin{cases} M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) &= c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \\ \mathbf{i}_{\mathrm{E}}^{[k]} &= f_{\mathrm{E0}}(\mathbf{u}^{[k]}), \end{cases} \tag{6.40a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA}}(\mathbf{i}_{\mathrm{E}}^{[k-1]}), \tag{6.40b}$$

$$\begin{cases} M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) &= c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \\ \mathbf{i}_{\mathrm{E}}^{[k]} &= f_{\mathrm{E1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}), \end{cases} \tag{6.41a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA}}(\mathbf{i}_{\mathrm{E}}^{[k-1]}), \tag{6.41b}$$

$$\begin{cases} M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) &= c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}), \\ \mathbf{i}_{\mathrm{E}}^{[k]} &= f_{\mathrm{E2}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}), \end{cases} \tag{6.42a}$$

$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA}}(\mathbf{i}_{\mathrm{E}}^{[k-1]}) \tag{6.42b}$$

Again, we figure that some of these Jacobi type WR schemes are analytically equivalent. In fact there are two equivalence classes as for the Gauss-Seidel type and these classes contain the same versions.

Lemma 6.16
Given the common initial guesses $\mathbf{x}^{[0]}$ *and* $\mathbf{u}^{[0]}$*, then, for any* $k \in \mathbb{N}$*, we obtain the following equivalence statements:*

(i) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *is a solution of either* (6.36), (6.37), (6.38) *and* (6.39) *or none of them.*

(ii) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]})$ *is a solution of* (6.40) *if and only if* $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *is a solution of* (6.35) *and*

$$\mathbf{i}_{\mathrm{E}}^{[k]} = f_{\mathrm{E0}}(\mathbf{u}^{[k]});$$

(iii) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]})$ *is a solution of* (6.41) *if and only if* $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *is a solution of* (6.35) *and*

$$\mathbf{i}_{\mathrm{E}}^{[k]} = f_{\mathrm{E1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{x}^{[k-1]});$$

(iv) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]})$ *is a solution of* (6.42) *if and only if* $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *is a solution of* (6.35) *and*

$$\mathbf{i}_{\mathrm{E}}^{[k]} = f_{\mathrm{E2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{x}^{[k-1]}).$$

Proof. The statement (i) if proven analogously by a conglomerate of the proofs of Lemmata 6.10 and 6.12. (ii), (iii) and (iv) is similarly shown as 6.14. □

Corollary 6.17
There are two equivalence classes of Jacobi type waveform relaxation scheme applied to the field/circuit coupled systems among the just introduced schemes.

Jac1 *The first class consists of the incorporated scheme* (6.35) *and all black-box schemes* (6.40), (6.41) *and* (6.42).

Jac2 *The second class contains the incorporated schemes* (6.36) *and* (6.37) *as well as all shifted variants* (6.38) *and* (6.39).

The classes representatives, for the convergence analysis, will be (6.35) for Jac1 and (6.39) for Jac2.

6.3.3 Convergence Analysis

As the convergence theorems in [PT18; JK96; Sau19] do not fit forms of inherent ODEs, which we obtain after decoupling of the full-Maxwell field/circuit coupled problems, we first provide a theorem that does. To be more precise, we need a convergence theorem, that considers iterates and their derivatives up to two steps before, see (6.60), (A.7), (A.12) and (A.15).

A Convergence Theorem

Let $\mathcal{I} = [t_0, T] \subset \mathbb{R}$ be a compact time interval and $\|\cdot\|$ a vector norm in $\mathbb{R}^n$. Now, for $t \in \mathcal{I}$, consider the following implicit ODE systems with initial condition

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(\frac{\mathrm{d}}{\mathrm{d}t}x(t), x(t), t) \ , \quad x(t_0) = x_0 \ , \tag{6.43}$$

where $f : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is a continuous function and and $x_0 \in \mathbb{R}^n$. Anyhow, contrary to the standard representation of WR methods, such as in Section 6.3, we now pay attention to one that is also dependent on the penultimate iterative solution, represented by

$$\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}(t) = F(\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}, \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}, \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-2]}, x^{[k]}, x^{[k-1]}, x^{[k-2]}, t) \ , \quad x^{[k]}(t_0) = x_0 \ , \tag{6.44}$$

for $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$, iteration counter $k > 2$ and initial guesses $x^{[0]}, x^{[1]} \in \mathbb{R}^n$, so that $x^{[0]}(t_0) = x^{[1]}(t_0) = x_0$, and complying

$$F(y, y, y, x, x, x, t) = f(y, x, t).$$

Assumption 6.18
Let F be Lipschitz continuous in the first six arguments introducing Lipschitz constants $c_1, c_2, c_3, d_1, d_2, d_3$ (see Definition A.10), and be continuous in all arguments. Further it holds $\sum_{i=1}^{3} c_i < 1$.

Theorem 6.19
Given Assumption 6.18, the initial value problem (6.43) has a unique solution $x^* \in C^1(\mathcal{I}, \mathbb{R}^n)$ to which the sequence $x^{[k]}$ of iterated solutions in $C^1(\mathcal{I}, \mathbb{R}^n)$ of WR scheme (6.44), converges in terms of C^1-norm as $k \to \infty$.

The proof of (6.19) will be given in the next section. Note that actually it is sufficient for F to be only Lipschitz continuous in the subsets of $\mathbb{R}^n$ where the iterated solutions' images live, see [Lel82]. The statement of Theorem 6.19 also applies to WR methods that rely only on one previous iterate, i. e

$$\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}(t) = G(\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}, \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}, x^{[k]}, x^{[k-1]}, t) \ , \quad x^{[k]}(t_0) = x_0 \ , \tag{6.45}$$

for $G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ complying $G(y, y, x, x, t) = f(y, x, t)$ with iteration counter $k > 1$ and initial guess $x^{[0]} \in \mathbb{R}^n$, so that $x^{[0]}(t_0) = x_0$. As already mentioned, this result is not new but can be verified as a special case of Theorem 6.19.

Assumption 6.20
Let G be Lipschitz continuous in the first four arguments introducing Lipschitz constants c_1, c_2, d_1, d_2 (see Definition A.10), and be continuous in all arguments. Further it holds $c_1 + c_2 < 1$.

Theorem 6.21
Given Assumption 6.20, then the initial value problem (6.43) has a unique solution $x^* \in C^1(\mathcal{I}, \mathbb{R}^n)$ to which the sequence $x^{[k]}$ of iterated solutions in $C^1(\mathcal{I}, \mathbb{R}^n)$ of WR scheme (6.45), converges in terms of C^1-norm as $k \to \infty$.

Proof. This becomes clear when defining F by

$$F(u_1, u_2, u_3, u_4, u_5, u_6, t) := G(u_1, u_2, u_4, u_5, t)$$

and choosing $x^{[1]}$ as the iterative solution of the scheme using G, for $k = 1$, since the choice of $x^{[1]}$ is not of importance for Theorem 6.19. Further, Assumption 6.18 is fulfilled with the same Lipschitz constants and additionally $c_3 = d_3 = 0$. Application of Theorem 6.21 proofs the statement. □

Concerning the convergence speed, the proof of Theorem 6.19 gives us some insight, see Remark 6.27.

Proof of Theorem 6.19

In this section we are going to proof Theorem 6.19 which relies on Nemytskii-type operators, cf. [Zei13; AE01; JK96]. Therefore, we start with some preliminaries.

Definition 6.22
We denote with $C^p(\mathcal{I}, \mathbb{R}^n)$ the Banach spaces of all p-times continuously differentiable functions from $\mathcal{I}$ to $\mathbb{R}^n$.

Definition 6.23
For any $p \in \mathbb{N}$, the C^p-norm of a function $f \in C^p(\mathcal{I}, \mathbb{R}^n)$ is defined by

$$\|f\|_{C^p} := \max_{0 \le q \le p} \left\| \frac{\mathrm{d}^q}{\mathrm{d}t^q} f \right\|_{C^0} ,$$

whereby

$$\|f\|_{C^0} := \sup_{t \in \mathcal{I}} \|f(t)\| .$$

Given WR scheme (6.44), we introduce the Nemytskii-type operator

$$\begin{aligned} \widetilde{F} :& C^0(\mathcal{I}, \mathbb{R}^n) \times C^0(\mathcal{I}, \mathbb{R}^n) \times C^0(\mathcal{I}, \mathbb{R}^n) \times C^1(\mathcal{I}, \mathbb{R}^n) \times C^1(\mathcal{I}, \mathbb{R}^n) \times C^1(\mathcal{I}, \mathbb{R}^n) \to C^0(\mathcal{I}, \mathbb{R}^n); \\ & (u_1, u_2, u_3, u_4, u_5, u_6) \mapsto \widetilde{F}(u_1, u_2, u_3, u_4, u_5, u_6) \end{aligned} \tag{6.46}$$

complying, for all $t \in \mathcal{I}$, the property

$$\widetilde{F}(u_1, u_2, u_3, u_4, u_5, u_6)(t) = F(u_1(t), u_2(t), u_3(t), u_4(t), u_5(t), u_6(t), t).$$

Then, we can rewrite (6.43) as

$$\frac{\mathrm{d}}{\mathrm{d}t} x(t) = \widetilde{F}(\frac{\mathrm{d}}{\mathrm{d}t} x, \frac{\mathrm{d}}{\mathrm{d}t} x, \frac{\mathrm{d}}{\mathrm{d}t} x, x, x, x)(t) , \quad x(t_0) = x_0 \tag{6.47}$$

and the WR scheme (6.44) as

$$\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}(t) = \widetilde{F}(\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}, \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}, \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-2]}, x^{[k]}, x^{[k-1]}, x^{[k-2]})(t)\ , \quad x^{[k]}(t_0) = x_0\ , \tag{6.48}$$

for iteration counter $k > 2$ and initial guesses $x^{[0]}, x^{[1]} \in C^1(\mathcal{I}, \mathbb{R}^n)$ so that $x^{[0]}(t_0) = x^{[1]}(t_0) = x_0$. It is easy to see that the solution $x \in C^1(\mathcal{I}, \mathbb{R}^n)$ to (6.47), if it exists, also represents a solution to (6.43). Moreover, the iterated solutions $x^{[k]} \in C^1(\mathcal{I}, \mathbb{R}^n)$ to (6.48) are also solutions to (6.44). As a consequence, if the sequence $x^{[k]}$ converges to a fixed point $x \in C^1(\mathcal{I}, \mathbb{R}^n)$, it solves the initial value problem (6.43).

Next, we define additional norms and provide a few Lemmata, cf. [Lel82; JK96].

Definition 6.24
Given a function $f \in C^0(\mathcal{I}, \mathbb{R}^n)$, we define the time interval's length as $H := T - t_0$ and additional norms by

$$\|f\|_h := \sup_{0 \le \bar{h} \le h} \left\|f(t_0 + \bar{h})\right\| \quad \text{and} \quad \|f\|_\lambda := \sup_{0 \le h \le H} \{\|f\|_h\, e^{-\lambda h}\}\ .$$

Note that for $h = H$ the norms $\|\cdot\|_h$ equals the C^0-norm, that is $\|\cdot\|_H = \|\cdot\|_{C^0}$.

Lemma 6.25
Consider the Nemytzkii operator $\widetilde{F}$ of (6.46) induced by an F fulfilling Assumption 6.18. Then there exist independent constants $c_i > 0$ and $d_i > 0$, for $i = 1, \dots, 3$, with $\sum_{i=1}^3 c_i < 1$ so that $\widetilde{F}$ complies the Lipschitz condition

$$\left\|\widetilde{F}(u_1, u_2, u_3, u_4, u_5, u_6) - \widetilde{F}(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4, \bar{u}_5, \bar{u}_6)\right\|_h \le \sum_{i=1}^{3} c_i \left\|u_i - \bar{u}_i\right\|_h + d_i \left\|u_{i+3} - \bar{u}_{i+3}\right\|_h\ ,$$

for all $u_1, u_2, u_3, \bar{u}_1, \bar{u}_2, \bar{u}_3 \in C^0(\mathcal{I}, \mathbb{R}^n)$, $u_4, u_5, u_6, \bar{u}_4, \bar{u}_5, \bar{u}_6 \in C^1(\mathcal{I}, \mathbb{R}^n)$ and $0 \le h \le H$.

Proof. The inequality is directly shown by

$$\begin{aligned}
&\left\|\widetilde{F}(u_1, u_2, u_3, u_4, u_6, u_6) - \widetilde{F}(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4, \bar{u}_5, \bar{u}_6)\right\|_h \\
&= \sup_{0 \le s \le h} \left\|\widetilde{F}(u_1, u_2, u_3, u_4, u_6, u_6)(t_0 + s) - \widetilde{F}(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4, \bar{u}_5, \bar{u}_6)(t_0 + s)\right\| \\
&= \sup_{0 \le s \le h} \Big\|F(u_1(t_0 + s), u_2(t_0 + s), u_3(t_0 + s), u_4(t_0 + s), u_5(t_0 + s), u_6(t_0 + s), t_0 + s) \\
&\qquad - F(\bar{u}_1(t_0 + s), \bar{u}_2(t_0 + s), \bar{u}_3(t_0 + s), \bar{u}_4(t_0 + s), \bar{u}_5(t_0 + s), \bar{u}_6(t_0 + s), t_0 + s)\Big\| \\
&\le \sup_{0 \le s \le h} \sum_{i=1}^{3} c_i \left\|u_i(t_0 + s) - \bar{u}_i(t_0 + s)\right\| + d_i \left\|u_{i+3}(t_0 + s) - \bar{u}_{i+3}(t_0 + s)\right\| \\
&\le \sum_{i=1}^{3} c_i \left\|u_i - \bar{u}_i\right\|_h + d_i \left\|u_{i+3} - \bar{u}_{i+3}\right\|_h .
\end{aligned}$$

□

Next, we recall a useful Lemma by Lelarasme [Lel82] in the variant of [JK96] for which, in accordance with x_0 of th IVP (6.43), we introduce the operator

$$(Jy)(t) := x_0 + \int_{t_0}^{t} y(s)\,\mathrm{d}s\ .$$

Note that for $y \in C^0(\mathcal{I}, \mathbb{R}^n)$, it is $Jy \in C^1(\mathcal{I}, \mathbb{R}^n)$ by the *fundamental theorem of calculus*.

Lemma 6.26
Let $y, \bar{y} \in C^0(\mathcal{I}, \mathbb{R}^n)$. Then, the following inequalities hold, for $0 \le h \le H$:

$$\|y - \bar{y}\|_h \le e^{\lambda h} \|y - \bar{y}\|_\lambda \qquad \textit{and} \qquad \|Jy - J\bar{y}\|_h \le \lambda^{-1} e^{\lambda h} \|y - \bar{y}\|_\lambda\ .$$

Proof. Using the norms' definition, the statement follows from

$$\|y - \bar{y}\|_h = e^{\lambda h} \|y - \bar{y}\|_h\, e^{-\lambda h} \le e^{\lambda h} \sup_{0 \le h \le H} \{\|y - \bar{y}\|_h\, e^{-\lambda h}\} = e^{\lambda h} \|y - \bar{y}\|_\lambda$$

and

$$\begin{aligned}
\|Jy - J\bar{y}\|_h &= \sup_{0 \le \bar{h} \le h} \left\| \int_{t_0}^{t_0 + \bar{h}} y(s) - \bar{y}(s)\,\mathrm{d}s \right\| = \sup_{0 \le \bar{h} \le h} \left\| \int_0^{\bar{h}} y(t_0 + s) - \bar{y}(t_0 + s)\,\mathrm{d}s \right\| \\
&\le \sup_{0 \le \bar{h} \le h} \int_0^{\bar{h}} \|y(t_0 + s) - \bar{y}(t_0 + s)\|\,\mathrm{d}s \le \int_0^h \|y - \bar{y}\|_s\,\mathrm{d}s \\
&\le \int_0^h e^{\lambda s} \|y - \bar{y}\|_\lambda\,\mathrm{d}s = \int_0^h e^{\lambda s}\,\mathrm{d}s\, \|y - \bar{y}\|_\lambda \le \lambda^{-1} e^{\lambda h} \|y - \bar{y}\|_\lambda\ .
\end{aligned}$$

□

Proof of Theorem 6.19. Looking at (6.47), we substitute $y(t) = \frac{\mathrm{d}}{\mathrm{d}t} x(t)$ and figure that $x(t) = (Jy)(t)$. Therefore, we can further rewrite (6.47) as

$$y(t) = \widetilde{F}(y, y, y, Jy, Jy, Jy)(t).$$

We define the mapping

$$\begin{aligned}
G : C^0(\mathcal{I}, \mathbb{R}^n) \times C^0(\mathcal{I}, \mathbb{R}^n) \times C^0(\mathcal{I}, \mathbb{R}^n) &\to C^0(\mathcal{I}, \mathbb{R}^n); \\
(y_1, y_2, y_3) &\mapsto G(y_1, y_2, y_3) := \widetilde{F}(y_1, y_2, y_3, Jy_1, Jy_2, Jy_3).
\end{aligned}$$

Now, for $i = 1, \ldots, 6$, $y_i, \bar{y}_i \in C^0(\mathcal{I}, \mathbb{R}^n)$ and all $t \in \mathcal{I}$, we make use of Lemmata 6.25 and 6.26

in order to bound

$$\begin{aligned}
&\|G(y_1,y_2,y_3)(t) - G(\bar{y}_1,\bar{y}_2,\bar{y}_3)(t)\| \\
&= \left\|\widetilde{F}(y_1,y_2,y_3,Jy_1,Jy_2,Jy_3)(t) - \widetilde{F}(\bar{y}_1,\bar{y}_2,\bar{y}_3,J\bar{y}_1,J\bar{y}_2,J\bar{y}_3)(t)\right\| \\
&\leq \sum_{i=1}^{3} c_i \|y_i - \bar{y}_i\|_h + d_i \|Jy_i - J\bar{y}_i\|_h \\
&\leq e^{\lambda h} \sum_{i=1}^{3} c_i \|y_i - \bar{y}_i\|_\lambda + d_i\lambda^{-1} \|y_i - \bar{y}_i\|_\lambda \\
&= e^{\lambda h} \sum_{i=1}^{3} (c_i + d_i\lambda^{-1}) \|y_i - \bar{y}_i\|_\lambda .
\end{aligned}$$

Since this bound is valid for all $t \in \mathcal{I}$, we obtain similarly

$$\|G(y_1,y_2,y_3) - G(\bar{y}_1,\bar{y}_2,\bar{y}_3)\|_\lambda \leq \sum_{i=1}^{3} (c_i + d_i\lambda^{-1}) \|y_i - \bar{y}_i\|_\lambda . \tag{6.49}$$

As of Lemma 6.25, it is $c_1 + c_2 + c_3 < 1$. Thus, we find $\varepsilon > 0$ with $c_1 + c_2 + c_3 + \varepsilon < 1$. Now, we can choose a $\lambda > 0$ so that $\lambda^{-1}(d_1 + d_2 + d_3) \leq \varepsilon$ yielding

$$\|G(y_1,y_2,y_3) - G(\bar{y}_1,\bar{y}_2,\bar{y}_3)\|_\lambda \leq (c_1 + c_2 + c_3 + \varepsilon) \max_{i=1,2,3} \{\|y_i - \bar{y}_i\|_\lambda\}.$$

Therefore, the mapping $T : C^0(\mathcal{I},\mathbb{R}^n) \to C^0(\mathcal{I},\mathbb{R}^n)$ with $T(y) := G(y,y,y)$ is a contraction mapping on $C^0(\mathcal{I},\mathbb{R}^n)$ and, by using the *Banach fixed point theorem*, admits a unique fixed-point $y^* \in C^0(\mathcal{I},\mathbb{R}^n)$ in terms of $\|\cdot\|_\lambda$, i. e. $y^* = T(y^*) = G(y^*,y^*,y^*)$. It follows, that $x^* := Jy^* \in C^1(\mathcal{I},\mathbb{R}^n)$ is the unique solution to (6.47) and thus to (6.43).
In the here introduced notation, the WR scheme represented by (6.48) translates, for $y^{[0]} = \frac{\mathrm{d}}{\mathrm{d}t}x^{[0]}$, to the sequence $y^{[k]}$ defined by

$$y^{[k]} = G(y^{[k]}, y^{[k-1]}, y^{[k-2]}).$$

Subtracting $y^{[k]}$ from the fixed-point y^* yields, in analogy to (6.49), the following inequality

$$\left\|y^* - y^{[k]}\right\|_\lambda = \left\|G(y^*,y^*,y^*) - G(y^{[k]},y^{[k-1]},y^{[k-2]})\right\|_\lambda \leq \sum_{i=1}^{3} (c_i + d_i\lambda^{-1}) \left\|y^* - y^{[k-i+1]}\right\|_\lambda$$

from which we conclude that

$$\left\|y^* - y^{[k]}\right\|_\lambda \leq \underbrace{\frac{c_2 + d_2\lambda^{-1}}{1 - (c_1 + d_1\lambda^{-1})}}_{=:e_1} \left\|y^* - y^{[k-1]}\right\|_\lambda + \underbrace{\frac{c_3 + d_3\lambda^{-1}}{1 - (c_1 + d_1\lambda^{-1})}}_{=:e_2} \left\|y^* - y^{[k-2]}\right\|_\lambda . \tag{6.50}$$

By defining the differences $v^{[k]} := \left\|y^* - y^{[k]}\right\|_\lambda$, (6.50) simplifies to

$$v^{[k]} \leq e_1 v^{[k-1]} + e_2 v^{[k-2]}.$$

In order to show that $v^{[k]}$ tends to zero as $k \to \infty$, consider $w^{[k]} := (v^{[k]}, v^{[k-1]})$. From (6.50) and the triviality that $v^{[k-1]} \leq v^{[k-1]}$ follows successively

$$w^{[k]} \leq \underbrace{\begin{bmatrix} e_1 & e_2 \\ 1 & 0 \end{bmatrix}}_{=:M} w^{[k-1]} \leq \begin{bmatrix} e_1 & e_2 \\ 1 & 0 \end{bmatrix}^2 w^{[k-2]} \leq \cdots \leq \begin{bmatrix} e_1 & e_2 \\ 1 & 0 \end{bmatrix}^{k-1} w^{[1]}. \tag{6.51}$$

Since e_1 and e_2 are positive numbers, for the eigenvalues of M, that is

$$\mu_{1,2} = \frac{e_1}{2} \pm \sqrt{\left(\frac{e_1}{2}\right)^2 + e_2},$$

holds $\mu_1 > 0 > \mu_2$ yielding the spectral radius

$$\eta(\lambda) := \mu_1 = \frac{c_2 + d_2\lambda^{-1}}{2 - 2(c_1 + d_1\lambda^{-1})} + \sqrt{\left(\frac{c_2 + d_2\lambda^{-1}}{2 - 2(c_1 + d_1\lambda^{-1})}\right)^2 + \frac{c_3 + d_3\lambda^{-1}}{1 - (c_1 + d_1\lambda^{-1})}}. \tag{6.52}$$

Now, $\eta(\lambda) < 1$ is equivalent to

$$\begin{aligned} & \frac{e_1}{2} + \sqrt{\frac{e_1}{4}^2 + e_2} < 1 \\ \Leftrightarrow \quad & \frac{e_1}{4}^2 + e_2 < \left(1 - \frac{e_1}{2}\right)^2 \\ \Leftrightarrow \quad & \frac{e_1}{4}^2 + e_2 < 1 - e_1 + \frac{e_1}{4}^2 \\ \Leftrightarrow \quad & e_2 < 1 - e_1. \end{aligned}$$

By definition, the latter expression reads

$$\begin{aligned} & \frac{c_3 + d_3\lambda^{-1}}{1 - (c_1 + d_1\lambda^{-1})} < 1 - \frac{c_2 + d_2\lambda^{-1}}{1 - (c_1 + d_1\lambda^{-1})} \\ \Leftrightarrow \quad & c_3 + d_3\lambda^{-1} < 1 - (c_1 + d_1\lambda^{-1}) - (c_2 + d_2\lambda^{-1}) \\ \Leftrightarrow \quad & c_1 + c_2 + c_3 + (d_1 + d_2 + d_3)\lambda^{-1} < 1 \end{aligned}$$

which is satisfied by the previous choice of λ. Consequently, $M^k \to 0$ as $k \to \infty$, which states that $w^{[k]} \to 0$ and thus $v^{[k]} \to 0$. To be more precise, we have linear convergence of $y^{[k]} \to y^*$ with rate $\eta(\lambda)$ in terms of norm $\|\cdot\|_\lambda$.
In order to show C^1-convergence of $x^{[k]}$, we first notice that we can express M in terms of a diagonal matrix containing the eigenvalues as follows

$$M = T \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix} T^{-1}, \quad \text{for } T = \begin{pmatrix} \mu_1 & \mu_2 \\ 1 & 1 \end{pmatrix}.$$

This allows us to rewrite (6.51) as

$$w^{[k]} \leq T \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix} T^{-1} \cdots T \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix} T^{-1} w^{[1]} = T \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}^{k-1} T^{-1} w^{[1]}$$

implicating $w^{[k]} \leq \|T\|_\infty \eta(\lambda)^{k-1} \left\|T^{-1}\right\|_\infty \left\|w^{[1]}\right\|_\infty$. With $c := \|T\|_\infty \left\|T^{-1}\right\|_\infty$ we deduce

$$v^{[k]} \leq c\eta(\lambda)^{k-1} \max\{v^{[1]}, v^{[0]}\}.$$

Finally, C^1-convergence of $x^{[k]}$ is shown by

$$\begin{aligned}
\left\|x^* - x^{[k]}\right\|_{C^1} &= \max\{\left\|Jy^* - Jy^{[k]}\right\|_{C^0}, \left\|y^* - y^{[k]}\right\|_{C^0}\} \\
&= \max\{\left\|Jy^* - Jy^{[k]}\right\|_{H}, \left\|y^* - y^{[k]}\right\|_{H}\} \\
&\leq \max\{\lambda^{-1}e^{\lambda H}, e^{\lambda H}\} \left\|y^* - y^{[k]}\right\|_{\lambda} \\
&\leq c \cdot \max\{\lambda^{-1}e^{\lambda H}, e^{\lambda H}\}\left(\left\|y^* - y^{[1]}\right\|_{\lambda} + \left\|y^* - y^{[0]}\right\|_{\lambda}\right)\eta(\lambda)^{k-1} \to 0 \qquad (6.53)
\end{aligned}$$

as k tends to infinity. □

Remark 6.27
Looking at the proof of Theorem 6.19 we obtain from (6.53) that the convergence rate of $x^{[k]}$ in terms of C^1-norm only asymptotically matches $\eta(\lambda)$, as defined in (6.52). Particularly, the convergence is especially restrained by the factor $\max\{\lambda^{-1}e^{\lambda H}, e^{\lambda H}\}$. Expecting λ to be large but fixed, $e^{\lambda H}$ grows exponentially with the window size H of the time interval $\mathcal{I}$. Therefore, the time interval has to be chosen sufficiently small according to λ, that is complying $e^{\lambda H} \approx 1$, in order to observe the convergence rate of $\eta(\lambda)$ right from the beginning.

6.3.3.1 Convergence of Gauss-Seidel Type GS1

In this section we analyze Gauss-Seidel method's convergence for the first class of the coupled circuit and EM device systems that is, by Corollary 6.15, GS1. As the representative for GS1 we choose (6.25). The convergence analysis builds upon finding the inherent ODE and the requirements allowing for applying Theorem 6.19.

In analogy to Lemma 5.6, we can interpret the EM device (E) as a mock element by setting $\mathbf{i}_{\mathrm{M}}^{[k]} = f_{\mathrm{EM0}}(\mathbf{u}^{[k]})$ and $A_{\mathrm{M}} = A_{\mathrm{E}}$. Hence, we will be able to decouple the MNA subsystem using Corollary 3.15 in the forthcoming. But first we start by inverting the EM device's subsystem (6.25a). As this is subsystem is shared by all iteration schemes in GS1, GS2, Jac1 and Jac2, we can reuse this scaling for the upcoming Theorems 6.29, 6.34, 6.36 and 6.39. Recalling the incorporated variant of the iteration schemes (6.25), i. e.

$$\begin{aligned}
M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) &= c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}) \\
f_{\mathrm{MNA}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{MNA}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) &= c_{\mathrm{MNA3}}(\mathbf{u}^{[k]}),
\end{aligned}$$

we scale the EM devices' subsystem by exploiting the nonsingularity of M_{E}, see Lemma 5.4, with its inverse

$$M_{\mathrm{E}}^{-1} = \begin{bmatrix} (\widetilde{S}M_\varepsilon G)^{-1} & 0 & 0 \\ 0 & 0 & I \\ -G(\widetilde{S}M_\varepsilon G)^{-1} & M_\varepsilon^{-1} & -M_\varepsilon^{-1}M_\sigma \end{bmatrix}.$$

Thus, we obtain the equivalent EM devices' subsystem

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + \widetilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) = \widetilde{c}_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}) \tag{6.54}$$

with, using the definitions in (5.8) and (5.9),

$$\begin{aligned}
\widetilde{b}_{\mathrm{E}}(\mathbf{u}) &:= M_{\mathrm{E}}^{-1}\begin{bmatrix} 0 & \widetilde{S}M_\zeta G M_\xi \widetilde{S} M_\zeta & 0 \\ M_\sigma G & \widetilde{C} M_\nu C & 0 \\ 0 & 0 & -I \end{bmatrix}\mathbf{u} \\
&= \begin{bmatrix} 0 & (\widetilde{S}M_\varepsilon G)^{-1}\widetilde{S}M_\zeta G M_\xi \widetilde{S} M_\zeta & 0 \\ 0 & 0 & -I \\ M_\varepsilon^{-1}M_\sigma G & -G(\widetilde{S}M_\varepsilon G)^{-1}\widetilde{S}M_\zeta G M_\xi \widetilde{S} M_\zeta + M_\varepsilon^{-1}\widetilde{C}M_\nu C & M_\varepsilon^{-1}M_\sigma \end{bmatrix}\mathbf{u} \\
\widetilde{c}_{\mathrm{E}}(\mathbf{w}_{\mathrm{E}}, \mathbf{x}) &:= M_{\mathrm{E}}^{-1}\begin{bmatrix} 0 \\ -I \\ 0 \end{bmatrix}\mathbf{w}_{\mathrm{E}} + M_{\mathrm{E}}^{-1}\begin{bmatrix} 0 \\ -M_\sigma G_\Gamma \Lambda A_{\mathrm{E}}^\top \\ 0 \end{bmatrix}\mathbf{e} \\
&= \begin{bmatrix} 0 \\ 0 \\ -M_\varepsilon^{-1} \end{bmatrix}\mathbf{w}_{\mathrm{E}} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -M_\varepsilon^{-1}M_\sigma G_\Gamma \Lambda A_{\mathrm{E}}^\top & 0 & 0 \end{bmatrix}\mathbf{x}
\end{aligned}$$

Lemma 6.28
Given the Assumptions required by Theorem 6.29, there always exists a vector norm $\|\cdot\|_*$ *so that for the induced matrix norm* $\left\|\bar{M}_2\right\|_* < 1$ *with* $\bar{M}_2$ *defined in* (6.61).

Proof. From the decoupling in Theorem 3.11 with modifications in Corollary 3.15 we obtain, using the same substitutions as in Theorem 6.29 and while dropping the arguments for the sake of simplicity,

$$\begin{aligned}
\partial_{\mathbf{w}_{\mathrm{E}}}\tilde{c}_{\mathrm{E}} &= \begin{bmatrix} 0 \\ 0 \\ -M_\varepsilon^{-1} \end{bmatrix}, & \partial_{\mathbf{x}} d_{\mathrm{E-MNA}} &= \begin{bmatrix} M_\varepsilon G_\Gamma \Lambda A_{\mathrm{E}}^\top & 0 & 0 \end{bmatrix}, \\
T_0 &= \begin{bmatrix} P_{\mathrm{C}}Q_{\mathrm{e}} & 0 \\ 0 & \bar{Q}_{\mathrm{L}} \\ 0 & 0 \end{bmatrix}, & T_2 &= \begin{bmatrix} Q_{\mathrm{C}}P_{\mathrm{V}} & Q_{\mathrm{C}}Q_{\mathrm{V}}P_{\mathrm{R}} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{P}_{\mathrm{V}} \end{bmatrix}, \\
\partial_{s_{cf}} f_2 &= \begin{bmatrix} 0 \\ \partial_{s_{cf}} f_{2r} \\ \partial_{s_{cf}} \bar{f}_{2v} \end{bmatrix}, & \partial_{\mathbf{y}} f_2 &= \begin{bmatrix} \partial_{\mathbf{y}} f_{2v} \\ \partial_{\mathbf{y}} f_{2r} \\ \partial_{\mathbf{y}} \bar{f}_{2v} \end{bmatrix} = \begin{bmatrix} \partial_{\mathbf{y}_e} f_{2v} & 0 \\ \partial_{\mathbf{y}_e} f_{2r} & \partial_{\bar{\mathbf{y}}_l} f_{2r} \\ \partial_{\mathbf{y}_e} \bar{f}_{2v} & \partial_{\bar{\mathbf{y}}_l} \bar{f}_{2v} \end{bmatrix}, \\
\partial_{\mathbf{u}} s_{cf} &= \begin{bmatrix} 0 & -A_{\mathrm{E}}\Lambda^\top G_\Gamma^\top \widetilde{C} M_\nu C & 0 \end{bmatrix}.
\end{aligned}$$

From this follows

$$\begin{aligned}
\frac{\partial \theta_1}{\partial \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} &= \partial_{\mathbf{w}_\mathrm{E}} \widetilde{c}_\mathrm{E} \partial_{\mathbf{x}} d_{\mathrm{E-MNA}} T_2 \partial_{s_{cf}} f_2 \partial_{\mathbf{u}} s_{cf} \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -G_\Gamma \Lambda A_\mathrm{E}^\top & 0 & 0 \end{bmatrix} \begin{bmatrix} Q_\mathrm{C} Q_\mathrm{V} P_\mathrm{R} \partial_{s_{cf}} f_{2r} \\ 0 \\ \bar{P}_\mathrm{V} \partial_{s_{cf}} \bar{f}_{2v} \end{bmatrix} \begin{bmatrix} 0 & -A_\mathrm{E} \Lambda^\top G_\Gamma^\top \widetilde{C} M_\nu C & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & G_\Gamma \Lambda A_\mathrm{E}^\top Q_\mathrm{C} Q_\mathrm{V} P_\mathrm{R} \partial_{s_{cf}} f_{2r} A_\mathrm{E} \Lambda^\top G_\Gamma^\top \widetilde{C} M_\nu C & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & U & 0 \end{bmatrix}, \\
\frac{\partial \theta_1}{\partial \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}} &= \partial_{\mathbf{w}_\mathrm{E}} \widetilde{c}_\mathrm{E} \partial_{\mathbf{x}} d_{\mathrm{E-MNA}} \left(T_0 + T_2 \partial_{\mathbf{y}} f_2 \right). \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -G_\Gamma \Lambda A_\mathrm{E}^\top & 0 & 0 \end{bmatrix} \left(\begin{bmatrix} P_\mathrm{C} Q_\mathrm{e} & 0 \\ 0 & \bar{Q}_\mathrm{L} \\ 0 & 0 \end{bmatrix} + \right. \\
&\quad \left. \begin{bmatrix} Q_\mathrm{C} P_\mathrm{V} \partial_{\mathbf{y}_e} f_{2v} + Q_\mathrm{C} Q_\mathrm{V} P_\mathrm{R} \partial_{\mathbf{y}_e} f_{2r} & Q_\mathrm{C} Q_\mathrm{V} P_\mathrm{R} \partial_{\bar{\mathbf{y}}_l} f_{2r} \\ 0 & 0 \\ \bar{P}_\mathrm{V} \partial_{\mathbf{y}_e} \bar{f}_{2v} & \bar{P}_\mathrm{V} \partial_{\bar{\mathbf{y}}_l} \bar{f}_{2v} \end{bmatrix} \right) \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ V & W \end{bmatrix}
\end{aligned}$$

with

$$U := G_\Gamma \Lambda A_\mathrm{E}^\top Q_\mathrm{C} Q_\mathrm{V} P_\mathrm{R} \partial_{s_{cf}} f_{2r} A_\mathrm{E} \Lambda^\top G_\Gamma^\top \widetilde{C} M_\nu C, \tag{6.55a}$$

$$V := -G_\Gamma \Lambda A_\mathrm{E}^\top \left(P_\mathrm{C} Q_\mathrm{e} + Q_\mathrm{C} P_\mathrm{V} \partial_{\mathbf{y}_e} f_{2v} + Q_\mathrm{C} Q_\mathrm{V} P_\mathrm{R} \partial_{\mathbf{y}_e} f_{2r} \right), \tag{6.55b}$$

$$W := -G_\Gamma \Lambda A_\mathrm{E}^\top \left(Q_\mathrm{C} Q_\mathrm{V} P_\mathrm{R} \partial_{\bar{\mathbf{y}}_l} f_{2r} \right). \tag{6.55c}$$

Consequently, $\bar{M}_2$ is a block matrix with zeros blocks for the block diagonals, i. e.

$$\bar{M}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & U & 0 & V & W \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

In order to find a vector norm $\|\cdot\|_*$ so that for the induced matrix norm holds $\left\|\bar{M}_2\right\|_* < 1$, we have to exploit the block matrix structure. Let $v = (v_1, v_2, v_3, v_4, v_5)$ be in the pre-image of $\bar{M}_2$. We choose a vector norm, e. g. the max-norm, and define

$$\varepsilon_U := \sup_{v_2 \neq 0} \frac{\|U v_2\|_\infty}{\|v_2\|_\infty}, \qquad \varepsilon_V := \sup_{v_4 \neq 0} \frac{\|V v_4\|_\infty}{\|v_4\|_\infty}, \qquad \varepsilon_W := \sup_{v_5 \neq 0} \frac{\|W v_5\|_\infty}{\|v_5\|_\infty}.$$

With these constants, we define the three custom norms

$$\|v_1\|_{2*} := 4\varepsilon_U \|v_2\|_\infty, \qquad \|v_1\|_{4*} := 4\varepsilon_V \|v_4\|_\infty, \qquad \|v_1\|_{5*} := 4\varepsilon_W \|v_5\|_\infty$$

and an overall norm for v by

$$\|v\|_* = \left\| \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} \right\|_* := \|v_1\|_\infty + \|v_2\|_{2*} + \|v_3\|_\infty + \|v_4\|_{4*} + \|v_5\|_{5*}\,. \tag{6.56}$$

For the induced matrix norm of $\|\cdot\|_*$ holds

$$\begin{aligned}
\left\|\bar{M}_2\right\|_* &= \sup_{v\neq 0} \frac{\left\|\bar{M}_2 v\right\|_*}{\|v\|_*} \\
&= \sup_{v\neq 0} \frac{\|Uv_2 + Vv_4 + Wv_5\|_\infty}{\|v_1\|_\infty + \|v_2\|_{2*} + \|v_3\|_\infty + \|v_4\|_{4*} + \|v_5\|_{5*}} \\
&\leq \sup_{v\neq 0} \frac{\|Uv_2\|_\infty + \|Vv_4\|_\infty + \|Wv_5\|_\infty}{\|v_2\|_{2*} + \|v_4\|_{4*} + \|v_5\|_{5*}} \\
&\leq \sup_{v\neq 0} \frac{\|Uv_2\|_\infty}{\|v_2\|_{2*}} + \sup_{v\neq 0} \frac{\|Vv_4\|_\infty}{\|v_4\|_{4*}} + \sup_{v\neq 0} \frac{\|Wv_5\|_\infty}{\|v_5\|_{5*}} \\
&= \sup_{v\neq 0} \frac{1}{4\varepsilon_U} \frac{\|Uv_2\|_\infty}{\|v_2\|_\infty} + \sup_{v\neq 0} \frac{1}{4\varepsilon_V} \frac{\|Vv_4\|_\infty}{\|v_4\|_\infty} + \sup_{v\neq 0} \frac{1}{4\varepsilon_W} \frac{\|Wv_5\|_\infty}{\|v_5\|_\infty} = \frac{3}{4} < 1.
\end{aligned}$$

□

Theorem 6.29 (convergence criterion for (6.25))
Let Assumptions 3.7, 3.8, 3.12, 4.20, 5.1, 5.2 and 3.10 be fulfilled. Further we assume that there is no LIM$^+$-cutset, that is Assumption 3.14. Then, the Gauss-Seidel type WR scheme's sequence $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ of (6.25) converges to the solution $(\mathbf{u}, \mathbf{x})$ of (5.21) in terms of C^0-norm on whole $\mathcal{I}$.

Proof. Since all the required assumptions are met, we can, for any k, decouple (6.25) by making use of Corollary 3.15 while setting $f := f_{\text{MNA1}}$, $d := d_{\text{MNA1}}$ and

$$s_c(\mathbf{i}_\text{M}) = A_\text{M}\mathbf{i}_\text{M} = A_\text{E} f_{\text{EM0}}(\mathbf{u}) =: s_{cf}(\mathbf{u})$$

yielding the equivalent system

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]} = f_0(\mathbf{y}^{[k]}, \mathbf{z}_1^{[k]}, \mathbf{z}_2^{[k]}, \mathbf{z}_3^{[k]}, s_i, s_{cf}(\mathbf{u}^{[k]})), \tag{6.57a}$$

$$\mathbf{z}_1^{[k]} = M_1(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}) \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k]} + f_1(\mathbf{y}^{[k]}, \mathbf{z}_2^{[k]}, \mathbf{z}_3^{[k]}, s_i, s_{cf}(\mathbf{u}^{[k]})), \tag{6.57b}$$

$$\mathbf{z}_2^{[k]} = f_2(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}, s_v, s_i, s_{cf}(\mathbf{u}^{[k]})), \tag{6.57c}$$

$$\mathbf{z}_3^{[k]} = M_3 \begin{pmatrix} s_i \\ s_v \end{pmatrix}. \tag{6.57d}$$

with the splitting $\mathbf{x}^{[k]} = T_0\mathbf{y}^{[k]} + T_1\mathbf{z}_1^{[k]} + T_2\mathbf{z}_2^{[k]} + T_3\mathbf{z}_3^{[k]}$ for transformation matrices

$$T_0 = \begin{bmatrix} P_\mathrm{C}Q_\mathrm{e} & 0 \\ 0 & \bar{Q}_\mathrm{L} \\ 0 & 0 \end{bmatrix},\ T_1 = \begin{bmatrix} Q_\mathrm{C}Q_\mathrm{V}Q_\mathrm{R} & 0 \\ 0 & 0 \\ 0 & \bar{Q}_\mathrm{V} \end{bmatrix},\ T_2 = \begin{bmatrix} Q_\mathrm{C}P_\mathrm{V} & Q_\mathrm{C}Q_\mathrm{V}P_\mathrm{R} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{P}_\mathrm{V} \end{bmatrix},\ T_3 = \begin{bmatrix} P_\mathrm{C}P_\mathrm{e} & 0 \\ 0 & \bar{P}_\mathrm{L} \\ 0 & 0 \end{bmatrix}.$$

Now, we resolve the variables in (6.57) by inserting them into each other

$$\begin{aligned}
\mathbf{z}_3^{[k]} &= M_3\begin{pmatrix} s_i \\ s_v \end{pmatrix} =: m_3(s_v, s_i) \\
\mathbf{z}_2^{[k]} &= f_2(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}, s_v, s_i, s_{cf}(\mathbf{u}^{[k]})) \\
&= f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i, s_{cf}(\mathbf{u}^{[k]})) \\
&=: f_2^{[k]} \\
\mathbf{z}_1^{[k]} &= M_1(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]})\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k]} + f_1(\mathbf{y}^{[k]}, \mathbf{z}_2^{[k]}, \mathbf{z}_3^{[k]}, s_i, s_{cf}(\mathbf{u}^{[k]})) \\
&= M_1(\mathbf{y}^{[k]}, m_3(s_v, s_i))m_3(\dot{s}_v, \dot{s}_i) + f_1(\mathbf{y}^{[k]}, f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i, s_{cf}(\mathbf{u}^{[k]})), \\
&\quad m_3(s_v, s_i), s_i, s_{cf}(\mathbf{u}^{[k]}))
\end{aligned}$$

and obtain for (6.57a)

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]} &= f_0(\mathbf{y}^{[k]}, \big[M_1(\mathbf{y}^{[k]}, m_3(s_v, s_i))\frac{\mathrm{d}}{\mathrm{d}t}m_3(s_v, s_i) \\
&\qquad + f_1(\mathbf{y}^{[k]}, f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i, s_{cf}(\mathbf{u}^{[k]})), m_3(s_v, s_i), s_i, s_{cf}(\mathbf{u}^{[k]}))\big], \\
&\qquad f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i, s_{cf}(\mathbf{u}^{[k]})), m_3(s_v, s_i), s_i, s_{cf}(\mathbf{u}^{[k]})) \\
&=: \theta_2(\mathbf{u}^{[k]}, \mathbf{y}^{[k]}, t)
\end{aligned} \tag{6.58}$$

Further, we have the derivatives

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k]} &= \frac{\mathrm{d}}{\mathrm{d}t}m_3(s_v, s_i), \\
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_2^{[k]} &= \frac{\partial}{\partial \mathbf{y}}f_2^{[k]}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]} + \frac{\partial}{\partial \mathbf{z}_3}f_2^{[k]}\frac{\mathrm{d}}{\mathrm{d}t}m_3(s_v, s_i) \\
&\qquad + \frac{\partial}{\partial s_v}f_2^{[k]}\frac{\mathrm{d}}{\mathrm{d}t}s_v + \frac{\partial}{\partial s_i}f_2^{[k]}\frac{\mathrm{d}}{\mathrm{d}t}s_i + \frac{\partial}{\partial s_{cf}}f_2^{[k]}\frac{\partial}{\partial \mathbf{u}}s_{cf}(\mathbf{u}^{[k]})\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}.
\end{aligned}$$

Due to Assumption 3.14, it is $A_\mathrm{E}^\top Q_\mathrm{C}Q_\mathrm{V}Q_\mathrm{R} = 0$ and, therefore, by definitions of linear $d_{\mathrm{E-MNA}}$ and c_E in (5.9), and thus $\widetilde{c}_\mathrm{E}$ in (6.54), we obtain

$$\begin{aligned}
d_{\mathrm{E-MNA}}(T_1 z_1) &= \begin{bmatrix} M_\varepsilon G_\Gamma \Lambda A_\mathrm{E}^\top & 0 & 0 \end{bmatrix}\begin{bmatrix} Q_\mathrm{C}Q_\mathrm{V}Q_\mathrm{R} & 0 \\ 0 & 0 \\ 0 & \bar{Q}_\mathrm{V} \end{bmatrix} z_1 = 0, \\
\widetilde{c}_\mathrm{E}(w, T_1 z_1) &= \begin{bmatrix} 0 \\ 0 \\ -M_\varepsilon^{-1} \end{bmatrix} w + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -M_\varepsilon^{-1}M_\sigma G_\Gamma \Lambda A_\mathrm{E}^\top & 0 & 0 \end{bmatrix}\begin{bmatrix} Q_\mathrm{C}Q_\mathrm{V}Q_\mathrm{R} & 0 \\ 0 & 0 \\ 0 & \bar{Q}_\mathrm{V} \end{bmatrix} z_1 = \widetilde{c}_\mathrm{E}(w, 0).
\end{aligned}$$

Now, we use the variable splitting $\mathbf{x}^{[k-1]} = T_0\mathbf{y}^{[k-1]} + T_1\mathbf{z}_1^{[k-1]} + T_2\mathbf{z}_2^{[k-1]} + T_3\mathbf{z}_3^{[k-1]}$, for (6.54), exploit the latter observations and insert the remaining algebraic expressions from the partial

decoupling. This allows us to rewrite (6.54), using the proper iteration counter, as

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} =& -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(T_0\mathbf{y}^{[k-1]} + T_1\mathbf{z}_1^{[k-1]} + T_2\mathbf{z}_2^{[k-1]} + T_3\mathbf{z}_3^{[k-1]}),\\
& T_0\mathbf{y}^{[k-1]} + T_1\mathbf{z}_1^{[k-1]} + T_2\mathbf{z}_2^{[k-1]} + T_3\mathbf{z}_3^{[k-1]})\\
=& -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(d_{\mathrm{E-MNA}}(T_0\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]} + T_2\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_2^{[k-1]} + T_3\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k-1]}),\\
& T_0\mathbf{y}^{[k-1]} + T_2\mathbf{z}_2^{[k-1]} + T_3\mathbf{z}_3^{[k-1]})\\
=& -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(d_{\mathrm{E-MNA}}(T_0\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]} + T_2\big[\frac{\partial}{\partial\mathbf{y}}f_2^{[k-1]}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}\\
& + \frac{\partial}{\partial\mathbf{z}_3}f_2^{[k-1]}\frac{\mathrm{d}}{\mathrm{d}t}m_3(s_v,s_i) + \frac{\partial}{\partial s_v}f_2^{[k-1]}\frac{\mathrm{d}}{\mathrm{d}t}s_v + \frac{\partial}{\partial s_i}f_2^{[k-1]}\frac{\mathrm{d}}{\mathrm{d}t}s_i\\
& + \frac{\partial}{\partial s_{cf}}f_2^{[k-1]}\frac{\partial}{\partial\mathbf{u}}s_{cf}(\mathbf{u}^{[k-1]})\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}\big] + T_3\frac{\mathrm{d}}{\mathrm{d}t}m_3(s_v,s_i)),\\
& T_0\mathbf{y}^{[k-1]} + T_2f_2^{[k-1]} + T_3m_3(s_v,s_i))\\
=:& \theta_1(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}, \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{u}^{[k-1]}, \mathbf{y}^{[k-1]}, t).
\end{aligned} \tag{6.59}$$

From (6.59) and (6.58) we obtain a WR scheme of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}(t) = \theta(\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}(t), \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}(t), x^{[k]}(t), x^{[k-1]}(t), t), \quad x^{[k]}(t_0) = (u_0, y_0) \tag{6.60}$$

with $\theta = (\theta_1, \theta_2)$, for $x^{[k]} = (\mathbf{u}^{[k]}, \mathbf{y}^{[k]})$, iteration counter $k \in \mathbb{N}$ and initial guess $x^{[0]} = (\cdot, \mathbf{y}^{[0]})$ so that $x^{[0]}(t_0) = (\cdot, y_0)$. Note that the previous dependency of t was dropped due to readability. Now, θ is continuous and Lipschitz continuous in the first four arguments by construction in Theorem 3.11, by Assumption 3.12 and by Lemma 5.5. Respectively, we denote the Lipschitz constants, which are independent of t, by c_1, c_2, d_1, d_2. In order to apply Theorem 6.21, by setting $G = \theta$, we first have to show that there exists a vector norm so that $c_1 + c_2 < 1$. Using the idea of the derivative test, see Lemma A.15, we consider the partial derivatives of θ with respect to $\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}$ and $\frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}$ yielding

$$\begin{aligned}
\bar{M}_1 &:= \frac{\partial\theta}{\partial\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}} = \begin{bmatrix} \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}} & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]}} \\ \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}} & \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]}} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},\\
\bar{M}_2 &:= \frac{\partial\theta}{\partial\frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}} = \begin{bmatrix} \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}} \\ \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} & \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}} \end{bmatrix} = \begin{bmatrix} \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}} \\ 0 & 0 \end{bmatrix}
\end{aligned} \tag{6.61}$$

with

$$\begin{aligned}
\frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} &= \partial_{\mathbf{w}_{\mathrm{E}}}\tilde{c}_{\mathrm{E}}\partial_{\mathbf{x}}d_{\mathrm{E-MNA}}T_2\partial_{s_{cf}}f_2\partial_{\mathbf{u}}s_{cf},\\
\frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}} &= \partial_{\mathbf{w}_{\mathrm{E}}}\tilde{c}_{\mathrm{E}}\partial_{\mathbf{x}}d_{\mathrm{E-MNA}}\left[T_0 + T_2\partial_{\mathbf{y}}f_2\right].
\end{aligned}$$

Note that especially due to the Lipschitz continuity of θ in the first two arguments, these partial derivatives are bounded. As of Lemma 6.28, there exists a vector norm $\|\cdot\|$ so that the induced matrix norm yields $\left\|\bar{M}_2\right\| < 1$. Since $\bar{M}_1 = 0$, it follows $\left\|\bar{M}_1\right\| + \left\|\bar{M}_2\right\| < 1$. Hence, we can apply Lemma A.15 twice guaranteeing that $c_1 + c_2 < 1$. From Theorem 6.21 we deduce that the sequence $x^{[k]}$ in (6.60) converges to $x^* = (\mathbf{u}^*, \mathbf{y}^*) \in C^1(\mathcal{I}, \mathbb{R}^n)$ which is the unique solution of

$$\frac{\mathrm{d}}{\mathrm{d}t}x = \theta(\frac{\mathrm{d}}{\mathrm{d}t}x, \frac{\mathrm{d}}{\mathrm{d}t}x, x, x, t), \quad x(t_0) = (u_0, y_0).$$

Then, we exploit the (Lipschitz) continuities of f_1, f_2 as well as the boundedness of M_1 and obtain for the limits of (6.57b)-(6.57d), as we assumed sufficient smooth source functions,

$$\begin{aligned}
\lim_{k\to\infty} \mathbf{z}_1^{[k]} &= \lim_{k\to\infty} M_1(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}) \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k]} + f_1(\mathbf{y}^{[k]}, \mathbf{z}_2^{[k]}, \mathbf{z}_3^{[k]}, s_i, s_{cf}(\mathbf{u}^{[k]})) \\
&= M_1(\mathbf{y}^*, \mathbf{z}_3^*) \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^* + f_1(\mathbf{y}^*, \mathbf{z}_2^*, \mathbf{z}_3^*, s_i, s_{cf}(\mathbf{u}^*)) = \mathbf{z}_1^*, \\
\lim_{k\to\infty} \mathbf{z}_2^{[k]} &= \lim_{k\to\infty} f_2(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}, s_v, s_i, s_{cf}(\mathbf{u}^{[k]})) \\
&= f_2(\mathbf{y}^*, \mathbf{z}_3^*, s_v, s_i, s_{cf}(\mathbf{u}^*)) = \mathbf{z}_2^*
\end{aligned}$$

where $\mathbf{z}_3^{[k]} = m_3(s_v, s_i) = \mathbf{z}_3^*$. Consequently,

$$\lim_{k\to\infty} \mathbf{x}^{[k]} = \lim_{k\to\infty} T_0\mathbf{y}^{[k]} + T_1\mathbf{z}_1^{[k]} + T_2\mathbf{z}_2^{[k]} + T_3\mathbf{z}_3^{[k]} = T_0\mathbf{y}^* + T_1\mathbf{z}_1^* + T_2\mathbf{z}_2^* + T_3\mathbf{z}_3^* = \mathbf{x}^*$$

converges in $C^0(\mathcal{I}, \mathbb{R}^n)$ with the according C^0-norm. □

Remark 6.30
If, for instance, Gauss-Seidel is used, the order of subsystems has an impact on the contraction constant and therefore it is of advantage to choose an order which minimizes it to obtain faster convergence, see [AG01].

With the help of Theorem 6.29 we found a first convergence criterion, that is the requirement of $\left\|\bar{M}_2\right\| < 1$. Note that this criterion is sufficient and not necessary. A further check on this criterion reveals that it is actually fulfilled without any additional requirements as shown in Lemma 6.28. In other words, it is the topological absence of LIE^+-cutset that guarantees the convergence. Additionally, when considering Corollary 6.15, this convergence result applies to all schemes within the first equivalence class of Gauss-Seidel type waveform iteration schemes which are (6.25), (6.31), (6.32) and (6.33) as of Corollary 6.15. We conclude first main convergence analysis for the coupled field / circuit systems by the following Corollary.

Corollary 6.31
Let Assumptions 3.7, 3.8, 3.12, 4.20, 5.1, 5.2 and 3.10 be fulfilled. If there is no LIE^+-cutset present, then the Gauss-Seidel type WR scheme's sequence

(i) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.25) *converges to the solution* $(\mathbf{u}, \mathbf{x})$ *of* (5.21)*;*

(ii) $(\mathbf{u}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.31) *converges to the solution* $(\mathbf{u}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x})$ *of* (5.26)*;*

(iii) $(\mathbf{u}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.32) *converges to the solution* $(\mathbf{u}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x})$ *of* (5.27)*;*

(iv) $(\mathbf{u}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.33) *converges to the solution* $(\mathbf{u}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x})$ *of* (5.28).

Proof. The absence of LIE^+-cutsets implies that Assumption 3.14 is fulfilled by Lemma 5.6. With no LIM^+-cutset present, all the requirements for Theorem 6.29 are met. As of Lemma 6.28, it is $\left\|\bar{M}_2\right\| < 1$ for an induced matrix norm. Thus, by Theorem 6.29 the sequence $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ of (6.25) converges to the solution $(\mathbf{u}, \mathbf{x})$ of (5.21). Since the sequences (6.31), (6.32) and (6.33) share the initial guess $\mathbf{x}^{[0]}$ with (6.25), Lemma 6.14 states that $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ together with respectively

$$\begin{aligned}
\mathbf{i}_{\mathrm{E}}^{[k]} &= f_{\mathrm{EM0}}(\mathbf{u}^{[k]}),\\
\mathbf{i}_{\mathrm{E}}^{[k]} &= f_{\mathrm{EM1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{x}^{[k-1]}) \quad \text{and}\\
\mathbf{i}_{\mathrm{E}}^{[k]} &= f_{\mathrm{EM2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{x}^{[k-1]})
\end{aligned}$$

solves (6.31), (6.32) and (6.33). Using the continuity properties, the limit $(\mathbf{u}^*, \mathbf{x}^*)$ imposes respectively

$$\begin{aligned}
\mathbf{i}_{\mathrm{E}}^{*} &= \lim_{k\to\infty} f_{\mathrm{EM0}}(\mathbf{u}^{[k]}) = f_{\mathrm{EM0}}(\mathbf{u}^{*}),\\
\mathbf{i}_{\mathrm{E}}^{*} &= \lim_{k\to\infty} f_{\mathrm{EM1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{x}^{[k-1]}) = f_{\mathrm{EM1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{*}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{*}, \mathbf{u}^{*}, \mathbf{x}^{*}),\\
\mathbf{i}_{\mathrm{E}}^{*} &= \lim_{k\to\infty} f_{\mathrm{EM2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{[k]}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{x}^{[k-1]}) = f_{\mathrm{EM2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{u}^{*}), \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^{*}, \mathbf{u}^{*}, \mathbf{x}^{*})
\end{aligned}$$

solving (5.26), (5.27) and (5.28). □

6.3.3.2 Convergence of Gauss-Seidel Type GS2

In this section we consider the second class of Gauss-Seidel type WR schemes for the coupled circuit and EM field device systems GS2. This time (6.30) will be the equivalence class' representative. Recalling this scheme, we have

$$\begin{aligned}
M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) &= c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]})\\
f_{\mathrm{MNA2}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA2}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) &= c_{\mathrm{MNA2.2}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}).
\end{aligned}$$

Again, the convergence analysis builds upon deducing an expression that allows the application of Theorem 6.19 given certain circumstances.

According to Remark 5.9, we can interpret the EM device as a mock element by setting $\mathbf{i}_{\mathrm{M}} = s_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u})$ and $A_{\mathrm{M}} = A_{\mathrm{E}}$. Further, we can also decouple the MNA subsystem (6.30b) as of Corollary 3.17. Note that the EM devices' subsystem (6.30a) is the same as (6.25a) so that it can also be rewritten as (6.54).

In order to guarantee convergence of GS2 we require an additional assumption.

Assumption 6.32
All EM field device branches are part of a V-loop.

Lemma 6.33
Given the matrices from Theorem 3.11 and mock elements with branches parallel to capacitors, i. e. A_M is a submatrix of A_C. Then, $Q_\mathrm{e}^\top P_\mathrm{C}^\top A_\mathrm{M} = 0$ if all mock element branches are part of a V-loop.

Proof. Without loss of generality we can write $A_\mathrm{C} = \begin{bmatrix} A_{\bar{c}} & A_\mathrm{M} \end{bmatrix}$ and from $Q_\mathrm{C}^\top A_\mathrm{C} = 0$ follows especially $Q_\mathrm{C}^\top A_\mathrm{M} = 0$.
Next, we figure that for each mock element branch a, that is a column of A_M, holds the loop property $\ker \begin{bmatrix} a & A_\mathrm{V} \end{bmatrix} \neq \{0\}$, according to Lemma A.6. Hence, $a \in \operatorname{im} A_\mathrm{V}$, since we do not allow V-loops, and therefore it exists an x fulfilling the representation $a = A_\mathrm{V} x$. Together with the first deduction, we obtain $0 = Q_\mathrm{C}^\top a = Q_\mathrm{C}^\top A_\mathrm{V} x$ stating that $x \in \ker Q_\mathrm{C}^\top A_\mathrm{V}$.
Further, from $\operatorname{im} \bar{Q}_\mathrm{V} = \ker P_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{V} = \ker Q_\mathrm{C}^\top A_\mathrm{V}$, where the latter is an immediate consequence of $Q_\mathrm{V}^\top Q_\mathrm{C}^\top A_\mathrm{V} = 0$, we follow that $x \in \operatorname{im} \bar{Q}_\mathrm{V}$. Finally, by definition it is $\operatorname{im} Q_\mathrm{e} = \ker \bar{Q}_\mathrm{V}^\top A_\mathrm{V}^\top P_\mathrm{C}$ which means that also $Q_\mathrm{e}^\top P_\mathrm{C}^\top A_\mathrm{V} \bar{Q}_\mathrm{V} = 0$ and since $x \in \operatorname{im} \bar{Q}_\mathrm{V}$,

$$0 = Q_\mathrm{e}^\top P_\mathrm{C}^\top A_\mathrm{V} x = Q_\mathrm{e}^\top P_\mathrm{C}^\top a.$$

As this holds for all columns a of A_M, we conclude $Q_\mathrm{e}^\top P_\mathrm{C}^\top A_\mathrm{M} = 0$. □

Theorem 6.34 (convergence criterion for (6.30))
Let Assumptions 3.7, 3.8, 3.12, 4.20, 5.1 and 5.2 be fulfilled. Further, we assume that every EM branch is part of a V-loop, that is Assumption 6.32. Then, the Gauss-Seidel type WR scheme's sequence $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ of (6.30) converges to the solution $(\mathbf{u}, \mathbf{x})$ of (5.25) in terms of C^0-norm on whole $\mathcal{I}$.

Proof. See Appendix A.4. □

Being part of a V-loop is a strong requirement for the EM decices' branches. Anyhow, this criterion is applicable to the whole second equivalence class of Gauss-Seidel WR schemes introduced in Corollary 6.15 which are (6.26), (6.27), (6.29) and (6.30). We conclude with the following Corollary.

Corollary 6.35
Let Assumptions 3.7, 3.8, 3.12, 4.20, 5.1 and 5.2 be fulfilled. If every EM branch is part of a V-loop (Assumption 6.32), then the Gauss-Seidel type WR scheme's sequence

(i) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.26) *converges to the solution* $(\mathbf{u}, \mathbf{x})$ *of* (5.22)*;*

(ii) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.27) *converges to the solution* $(\mathbf{u}, \mathbf{x})$ *of* (5.23)*;*

(iii) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.29) *converges to the solution* $(\mathbf{u}, \mathbf{x})$ *of* (5.24)*;*

(iv) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.30) *converges to the solution* $(\mathbf{u}, \mathbf{x})$ *of* (5.25)*.*

Proof. With the EM branches being part of a V-loops we can use Lemma A.16 and guarantee convergence as of Theorem 6.34. Since the sequences (6.26), (6.27) and (6.29) share the initial guess $\mathbf{x}^{[0]}$ with (6.30), Lemma 6.12 states that all sequences share the same solutions $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ for any $k \in \mathbb{R}$. This holds especially for the limit $(\mathbf{u}^*, \mathbf{x}^*)$ which, as a solution of (5.25), is then also a solution of (5.22), (5.23) and (5.25) by Corollary 6.15. □

6.3.3.3 Convergence of Jacobi Type Jac1

In this section we proceed entirely analogously to the previous sections but for the iteration scheme in class Jac1 of Corollary 6.17. Recalling the representative of Jac1, which is (6.35), we have

$$M_{\mathrm{E}} \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{u}^{[k]} + b_{\mathrm{E}}(\mathbf{u}^{[k]}) = c_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}^{[k-1]}), \mathbf{x}^{[k-1]}),$$
$$f_{\mathrm{MNA1}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{MNA1}}(\mathbf{x}^{[k]}), \mathbf{x}^{[k]}, t) = c_{\mathrm{MNA1.0}}(\mathbf{u}^{[k-1]}).$$

As the proof of the convergence behavior is similar to the ones of GS1, we start right away with the following statement:

Theorem 6.36 (convergence criterion for (6.35))
Let Assumptions 3.7, 3.8, 3.12, 4.20, 5.1, 5.2 and 3.10 be fulfilled. Further we assume that there is no LIM^+-cutset, that is Assumption 3.14. Then, the Jacobi type WR scheme's sequence $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ of (6.35) converges to the solution $(\mathbf{u}, \mathbf{x})$ of (5.21) in terms of C^0-norm on whole $\mathcal{I}$.

Proof. See Appendix A.4. □

Corollary 6.37
Let Assumptions 3.7, 3.8, 3.12, 4.20, 5.1, 5.2 and 3.10 be fulfilled. If there is no LIE^+-cutset present, then the Jacobi type WR scheme's sequence

(i) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.35) *converges to the solution* $(\mathbf{u}, \mathbf{x})$ *of* (5.21);

(ii) $(\mathbf{u}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.40) *converges to the solution* $(\mathbf{u}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x})$ *of* (5.26);

(iii) $(\mathbf{u}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.41) *converges to the solution* $(\mathbf{u}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x})$ *of* (5.27);

(iv) $(\mathbf{u}^{[k]}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.42) *converges to the solution* $(\mathbf{u}, \mathbf{i}_{\mathrm{E}}^{[k]}, \mathbf{x})$ *of* (5.28).

Proof. Analogue to Corollary 6.31 but with Lemma 6.16 □

Remark 6.38
Using Jacobi type WR schemes, the order of subsystems does not affect the contraction constant. In fact, the subsystems are entirely independent during an iteration sweep. This is why these schemes are predestined for parallelization.

6.3.3.4 Convergence of Jacobi Type Jac2

In this section we consider the second class of Jacobi type WR schemes for the coupled circuit and EM field device systems. The representative is (6.39).

Theorem 6.39 (Convergence criterion for (6.39))
Let Assumptions 3.7, 3.8, 3.12, 4.20, 5.1 and 5.2 be fulfilled. Further, we assume that every EM branch is part of a V-loop, that is Assumption 6.32. Then, the Jacobi type WR scheme's sequence $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.39) *converges to the solution* $(\mathbf{u}, \mathbf{x})$ *of* (5.25) *in terms of* C^0*-norm on whole* $\mathcal{I}$*.*

Proof. See Appendix A.4. □

We conclude with the following Corollary.

Corollary 6.40
Let Assumptions 3.7, 3.8, 3.12, 4.20, 5.1 and 5.2 be fulfilled. If every EM branch is part of a V-loop (Assumption 6.32), then the Jacobi type WR scheme's sequence

(i) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.36) *converges to the solution* $(\mathbf{u}, \mathbf{x})$ *of* (5.22)*;*

(ii) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.37) *converges to the solution* $(\mathbf{u}, \mathbf{x})$ *of* (5.23)*;*

(iii) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.38) *converges to the solution* $(\mathbf{u}, \mathbf{x})$ *of* (5.24)*;*

(iv) $(\mathbf{u}^{[k]}, \mathbf{x}^{[k]})$ *of* (6.39) *converges to the solution* $(\mathbf{u}, \mathbf{x})$ *of* (5.25)*.*

Proof. Analogue to Corollary 6.35 while using Corollary 6.17. □

6.4 Conclusions and Remarks

In this chapter we focused on the simulation of field/circuit coupled problems represented by systems introduced in Chapter 5. First, in Section 6.1, we provided the coupled modeling framework realized by the mock element interface. Then, in Section 6.2 we addressed a monolithic approach, where the systems are solved by numerical integration methods for DAEs such as backward differentiation formula or implicit Runge-Kutta methods. The resulting nonlinear system of equations are usually solved using Newton's method involving the solving of linear systems of equations. We discussed the direct and iterative approaches to deal with this kind of systems. In order to involve solving methods that are in particular efficient for subsystems, we introduced a hybrid method approach in Section 6.2.3 using the Schur complement. We further introduced scaling strategies to enhance the iterative solving process with respect to convergence and introduced a structural exploitation to significantly reduce the linear systems dimension which are supported by numerical benchmarks in Chapter 7.
Then, in Section 6.3 we studied the co-simulation approach with emphasis on waveform relaxation (WR) schemes. Here we introduced two popular types of WR schemes for coupled systems that have DAE form which are the Gauss-Seidel and Jacobi types. Contrary to the

analytical equivalence of the field/circuit coupled systems, when applying Gauss-Seidel type WR to them, they divide into two equivalence classes, denoted with GS1 and GS2 wherein the iterative solutions are equal, see Corollary 6.15. The same happens when applying Jacobi type WR methods to them, yielding Jac1 and Jac2 as of Corollary 6.17. In order to analyze GS1, GS2, Jac1 and Jac2 with regard to convergence, we first developed the convergence theorem, Theorem 6.19, in preparation for the main results on the equivalence classes. Finally, with the help of the decoupling theorems in Chapter 3, and Theorem 6.19 we were able to derive the convergence criterion in Theorems 6.29, 6.34, 6.36 and 6.39 which guarantee convergence of, respectively, GS1, GS2, Jac1 and Jac2 given certain topological assumptions. GS1, GS2, Jac1 and Jac2 are equivalence classes, for which we provided topological convergence criteria, see Corollaries 6.31, 6.35, 6.37 and 6.40. To give a few examples: If there is no LIE^{+}-cutset, (6.25) converges; if there is no voltage source parallel to the EM devices branches, (6.38) can not be guaranteed to converge.

Remarks on Monolithic and WR Approach An outstanding advantage of WR methods is that each subsystem can be simulated by tailored methods which are likely to perform much superior then non-specialized methods on the overall system. If these iteration schemes are guaranteed to converge, they are worth being considered since they allow each subsystem to act on a different scale that is appropriate for the wave propagation. If, however, the time scales are equal, a monolithic approach may be of interest especially when the simulator has access to the linear solving subroutine for each subsystem. In particular, using a hybrid method approach for the linear systems of equations one can bound the number of solving linear subsystems by Lemma 6.5 or Corollary 6.6 whereas the number of iterations required for WR schemes is unclear. Additionally, we know that a multirate approach, is not only possible for WR methods, they can also be simulated during a monolithic simulation process for instance by bypassing subsystem solutions until an update is required, see [SBD12]. Both variants still provide plenty possibilities of improvement, for instance acceleration techniques for WR methods, see e. g. [Cor20], or more structural exploitation for the monolithic approach. Moreover, parallelization is also possible for both approaches.

7 Numerical Benchmarks

In the following we provide benchmarks in support of the monolithic and waveform relaxation (WR) approach discussed in Chapter 6. Hereby, the latter is done in accordance with [ST20]. In order to deal with the exact model descriptions considered in this treatise, the benchmarks are produces by a self-implemented electric circuit simulator which is designed to communicate with both, a self-implemented version of an electromagnetic (EM) device modeler and `DevEM` via the same mock element based interface design introduced before. The non-black-box implementation is realized in `python` and uses the network structure in [Str+18] with automatic differentiation as a computational backed [GW08]. They communicate via the same interface originally developed for `DevEM` [Sch+16]. This enables the possibility of both, verifying the methods developed upon the herein considered model equations and investigating the behavior of industrial tools.

7.1 Monolithic Simulation Benchmarks

In the following we consider some examples of coupled field/circuit systems, that were simulated using the techniques introduced in Section 6.2. The first part, Section 7.1.1 considers the coupled systems presented in Chapter 5 and provides a proof of concept for the interface presented in Section 6.1 and the methods in Section 6.2. The second part goes further and shows that this monolithic concept also works for refined models from black-box modelers, which in this case is `DevEM`.

material	permittivity ε	conductivity σ	permeability μ
aluminum	$9.5 \cdot \varepsilon_0$	$3.33 \cdot 10^7$	μ_0
copper	ε_0	$5.7 \cdot 10^7$	μ_0
oxide	$3.9 \cdot \varepsilon_0$	0	μ_0
nickel-chrome	ε_0	891265.5971	μ_0

Table 7.1: Specific properties of commonly used materials.

7.1.1 Coupled Field/Circuit Systems

The first example of a coupled field/circuit is a simple low-pass filter.

Example 7.1 (low-pass filter with EM device)
Consider the coupled field/circuit problem in Figure 7.1

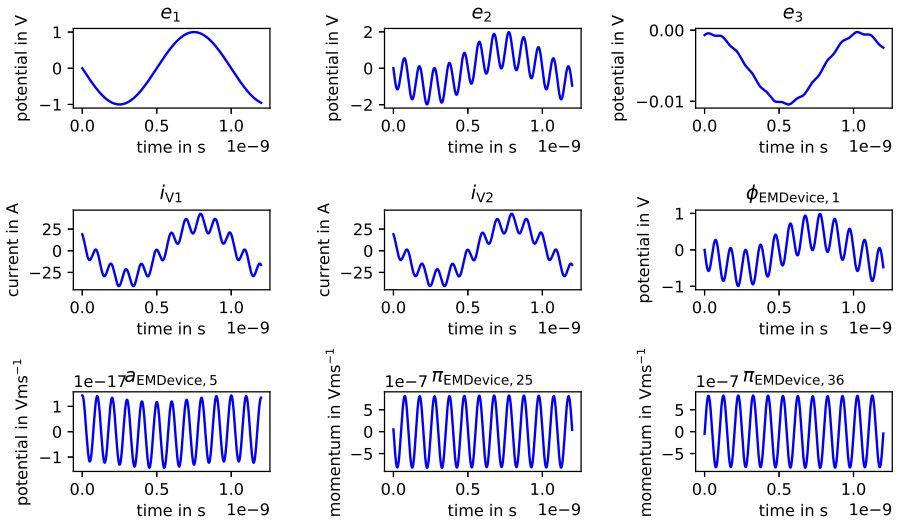


Figure 7.2: Simulation plots for low-pass filter with EM device of Example 7.1.

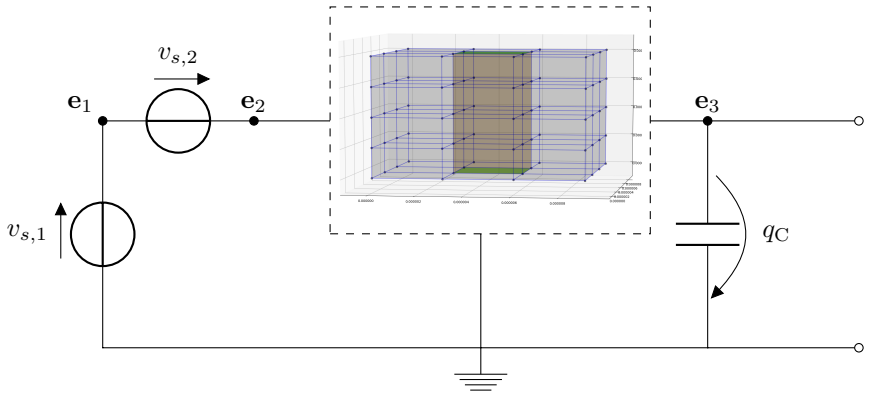


Figure 7.1: Low-pass filter using EM device as resistor with $v_{s,1}(t) = \sin(2\pi 10^9 t)$, $v_{s,2}(t) = \sin(2\pi 10^{10} t)$ and $q_C(t) = 10.307856 \times 10^{-07}$.

The quantities' dimensions are $(\mathbf{e}, \mathbf{i}_V, \boldsymbol{\phi}, \mathbf{a}, \boldsymbol{\pi}) : \mathcal{I} \to \mathbb{R}^{3+2+4+20+20}$.

Simulation of Example 7.1 using backward differentiation formula (BDF) of first order with constant time step size $h = 10^{-12}$ on $\mathcal{I} = [0, 1.2 \times 10^{-9}]$, Newton's method as nonlinear solver and `spsolve`[1] yields the results in Figure 7.2. Using `gmres`[2] the resulting linear systems of equations for Example 7.1 the following observations are made. As we can see from Table 7.2, we can advance the iterative solving process of the underlying linear system of equations by making use of the techniques introduce in Section 6.2. Concerning the simulation's error, the scaling and elimination techniques introduced in Section 6.2 have an impact in that they are

[1] direct solver for linear systems of equations from `pythons scipy` package.
[2] GMRES iterative solver for linear systems of equations from `pythons scipy` package.

row & var scaling	$\mathbf{a}/\boldsymbol{\pi}$-elimination	incomplete LU preconditioned	average iterations
x	x	x	fail
x	✓	x	174.51
x	x	✓	32.32
x	✓	✓	11.93
✓	x	✓	4.16
✓	✓	✓	3.04

Table 7.2: Monolithic simulation of Example 7.1 using GMRES for the linear solving routine.

row scaling	x	✓	x	✓	x	✓
var scaling	x	x	✓	✓	x	✓
$\mathbf{a}/\boldsymbol{\pi}$-elimination	x	x	x	x	✓	✓
$h = 10^{-10}$	44.33499	27.044333	36.998866	26.93026	26.952210	26.930268
$h = 10^{-11}$	6.674461	8.595887	5.684715	6.794140	6.634962	6.568160
$h = 10^{-12}$	0.654328	0.654048	0.645739	0.654035	0.654035	0.654035

Table 7.3: Errors for monolithic simulation of Example 7.1 using incomplete LU preconditioned `gmres`. Errors are given in infinity-norm.

slightly sensitive to the iterative method's tolerance of 10^{-10}, see Table 7.3. The reference solution was calculate with constant step size $h = 10^{-13}$ using direct solver. Contrary, to these observations, they do not have that much impact on direct solvers such as `spsolve` but for variable scaling alone as show in Table 7.4. Note that the $\mathbf{a}/\boldsymbol{\pi}$-elimination has a positive impact on direct solvers, such as `spsolve` or `scipy_lu`[3] due to the increased number of unknowns, see for instance in Figure 7.3 where the EM field device's mesh was refined to change the dimensions.

Consider again Example 7.1 where the mesh for the spartial discretization of the EM device is coarsened to $3 \times 3 \times 6$ volumes. At this point, solving the coupled field/circuit linear equation system during the simulation unsing `gmres` does not converge to a solution anymore if applied on the whole Jacobians, even with $\mathbf{a}/\boldsymbol{\pi}$-elimination. Contrary, if using a hybrid solving approach with `spsolve` applied to the circuit's subsystem's part of the Jacobian and `gmres` on the EM

[3] method of solving a linear system of equations using `scipy`'s LU decomposition

row scaling	x	x	✓	x	✓
var scaling	x	✓	✓	x	✓
$\mathbf{a}/\boldsymbol{\pi}$-elimination	x	x	x	✓	✓
$h = 10^{-10}$	26.93026	48.84513	26.93026	26.93026	26.93026
$h = 10^{-11}$	6.56816	6.71004	6.56816	6.56816	6.56816
$h = 10^{-12}$	0.65403	0.65403	0.65403	0.65403	0.65403

Table 7.4: Errors for monolithic simulation of Example 7.1 using `spsolve`. Errors are given in infinity-norm.

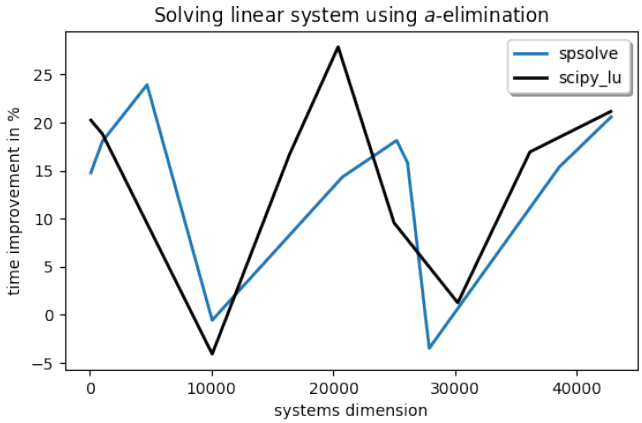


Figure 7.3: Performance gain in solving the linear systems due to $\mathbf{a}/\boldsymbol{\pi}$-elimination.

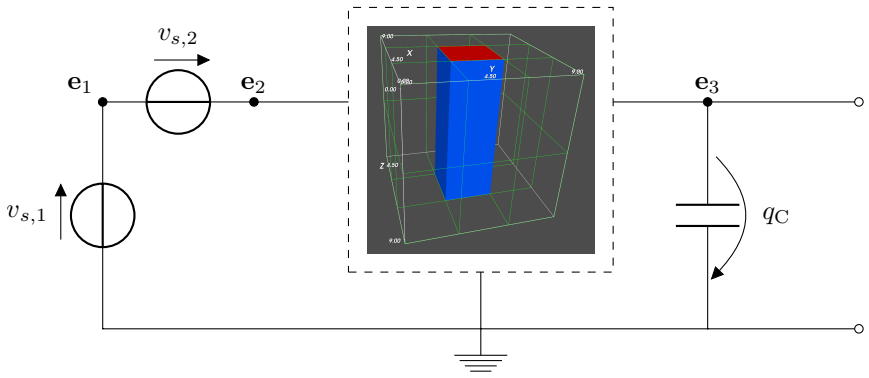


Figure 7.4: Low-pass filter using EM device from `DevEM` as resistor with $v_{s,1}(t) = \sin(2\pi 10^9 t)$, $v_{s,2}(t) = \sin(2\pi 10^{10} t)$ and $q_C(t) = 10.307856 \times 10^{-07}$.

device's one, the simulation succeeds with an average of 28,6198 iterations per linear solving step.

7.1.2 Using `DevEM` Models

As for testing the interface with third-party modelers, we consider the following 3D models as test cases for *MAGWEL*'s `DevEM` software package, see Table 7.5. The first test case is a most simple aluminum bar only contributing four discrete electric scalar potentials. Hence, an equation and variable-wise debugging can be performed. The same aluminum bar is also considered with enabled magnetic vector potentials.

Example 7.2 (low-pass filter with EM device from `DevEM`)
Consider the coupled field/circuit problem in Figure 7.4

The quantities' dimensions are $(\mathbf{e}, \mathbf{i}_V, \boldsymbol{\phi}, \mathbf{a}, \boldsymbol{\pi}) : \mathcal{I} \to \mathbb{R}^{3+2+4+8+8}$.

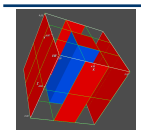
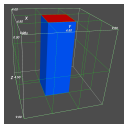
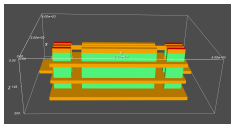
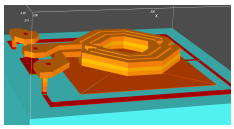
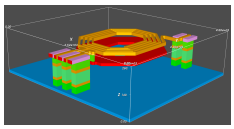
model	frequencies	system dimension	description
	1 MHz-10 GHz	7^2	aluminum bar
	1 MHz-10 GHz	82^2	aluminum bar
	10 GHz	255.427^2	*ACCO* transmission line
	1 GHz	493.547^2	*ACCO* inductor
	10 GHz	295.900^2	*ACCO* balun

Table 7.5: Refined model test cases for `DevEM`.

The geometry and material properties of the EM device in Example 7.2 match the ones of the EM device in Example 7.1. Thus, for the circuit in Figure 7.4 we obtain comparable simulation results to the ones in Figure 7.2, that is the ones on Figure 7.5. Note that the discrete magnetic vector and electric scalar potentials may be of different magnitude due to different gauge conditions used by `DevEM`.

Example 7.3 (balun circuit with `DevEM` device)
Consider the coupled field/circuit problem in Figure 7.6

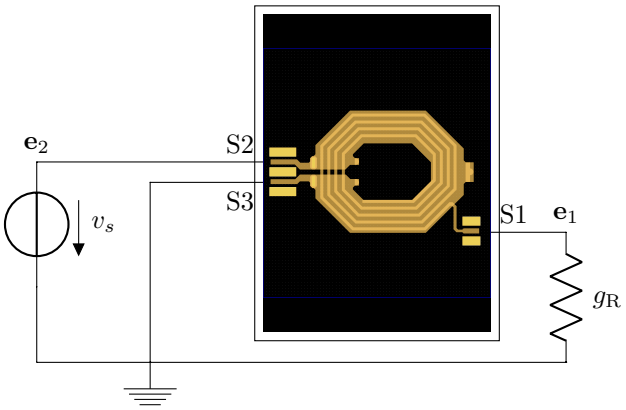


Figure 7.6: Balun test circuit using *ACCO*'s balun EM device from `DevEM` with $v_s(t) = \sin(2\pi 10^{10} t)$ and $g_\mathrm{R}(t) = 50$.

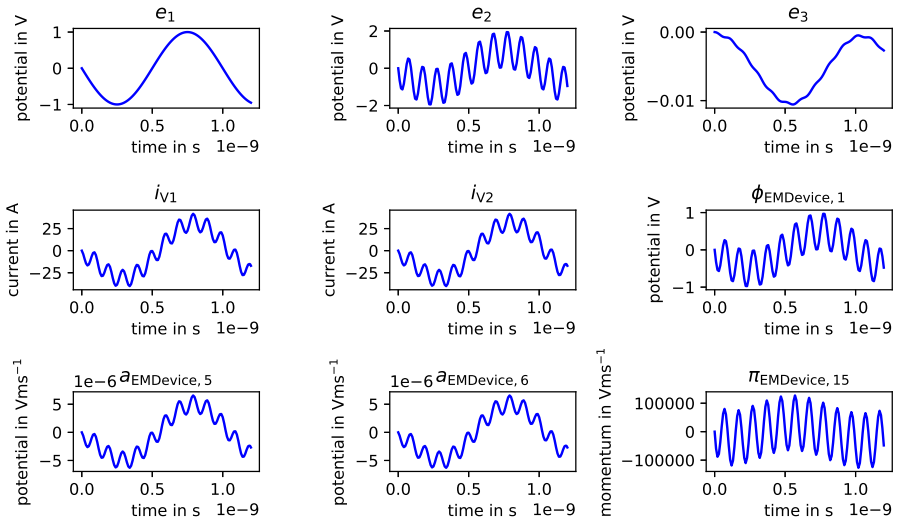


Figure 7.5: Simulation plots for low-pass filter with EM device from `DevEM` of Example 7.2.

The quantities' dimensions are $(\mathbf{e}, \mathbf{i}_V, \boldsymbol{\phi}, \mathbf{a}, \boldsymbol{\pi}) : \mathcal{I} \to \mathbb{R}^{2+1+46342+124775+124775}$.

The field/circuit problem in Example 7.3 was solved using $\mathbf{a}/\boldsymbol{\pi}$-elimination, variable and row scaling introduced in Section 6.2. Similar to Example 7.1 with a mesh consisting of $3 \times 3 \times 6$ volumes, not only `gmres` but also `spsolve` failed for the linear systems of equations. However, the hybrid approach was successful. To be more precise, we used the hybrid solving routine (6.14)-(6.17) whereby the circuit subsystem was solved using `spsolve` and the EM device subsystem using `MUMPS`[4], that is for the solving procedures (6.14), (6.15) and (6.17).

Remark 7.4
Solving the coupled field/circuit system solely with `MUMPS` as the liner solver was also successful. The hybrid approach, however, has proven to be successful and advantageous since it is not restricted to certain solvers which enables the use of tailored ones.

Results of Example 7.1 using BDF of first order with constant step size $h = 10^{-12}$ on $\mathcal{I} = [0, 10^{-10}]$, Newton's method as nonlinear solver are shown in Figure 7.7. Sectional magnetic flux density of the balun device for some sample time points are shown in Figure 7.8.

Also the simulation of multiple EM devices obtained from `DevEM` is possible. For this consider the band-pass filter of following example.

Example 7.5 (band-pass filter with `DevEM` devices)
Consider the coupled field/circuit problem with multiple refined models as obtained by `DevEM` in Figure 7.9

[4] MUltifrontal Massively Parallel Solver: A parallel sparse direct solver, see e. g. http://mumps-solver.org

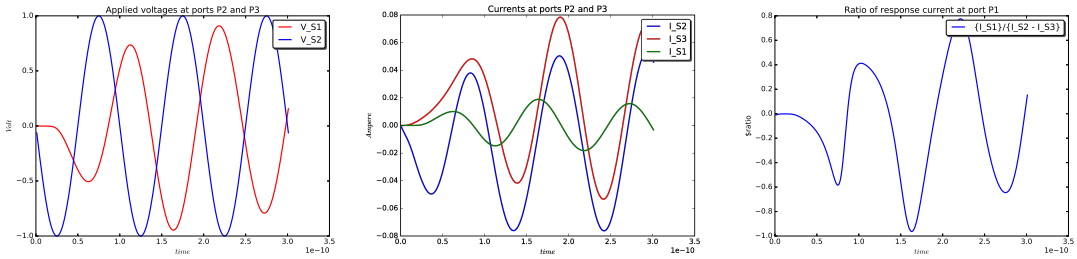
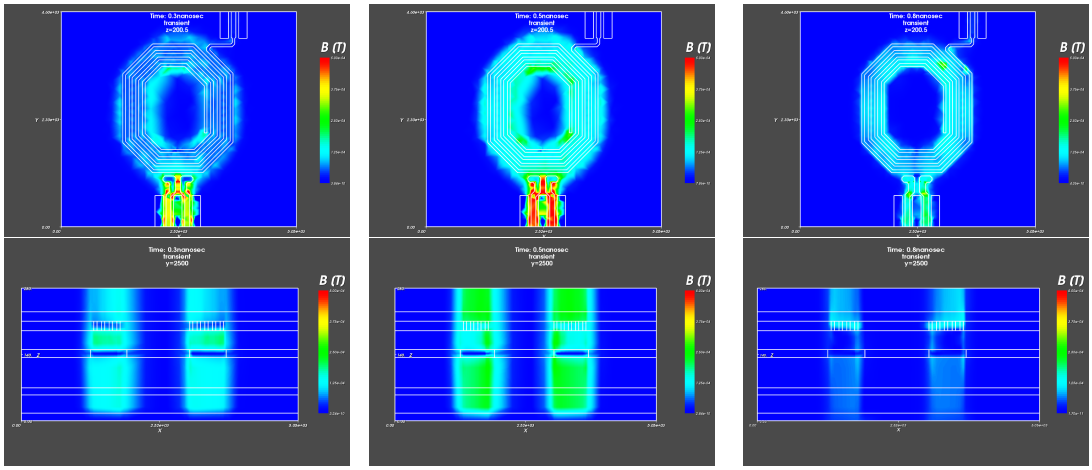


Figure 7.7: Simulation plots for balun circuit with `DevEM` device Example 7.3.



(a) $t = 0.3 \times 10^{-9}$ s **(b)** $t = 0.5 \times 10^{-9}$ s **(c)** $t = 0.8 \times 10^{-9}$ s

Figure 7.8: Sectional view of magnetic flux density of balun in Example 7.3.

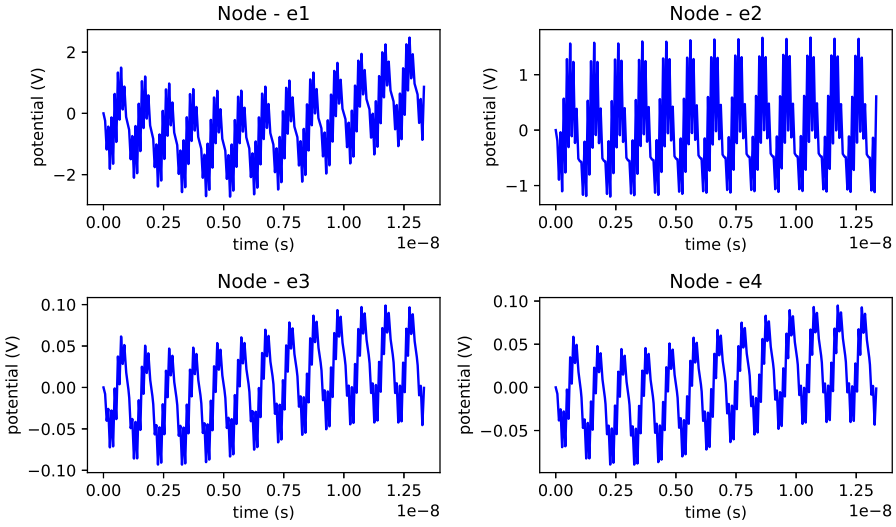


Figure 7.10: Simulation plots for band-pass filter with multiple EM devices of Example 7.5.

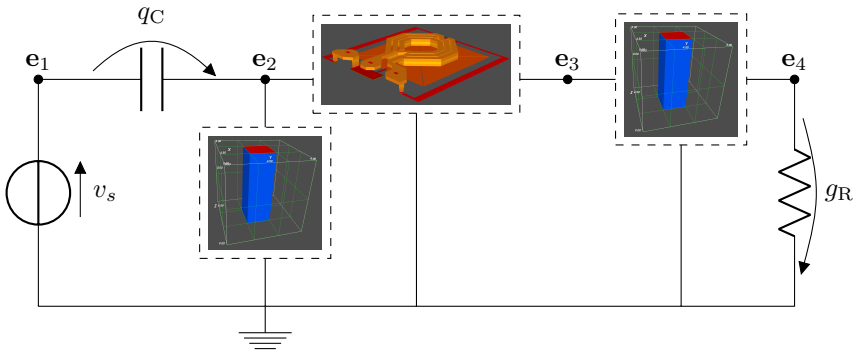


Figure 7.9: Band-pass filter using EM devices as resistors and inductor with $v_s(t) = \sin(\pi 10^8 t) + \sin(2\pi 10^9 t) + \sin(2\pi 0.8 \times 10^{10} t)$, $g_R(t) = 1$ and $q_C(t) = \frac{1}{2\pi 10^9}$.

The quantities' dimensions are $(\mathbf{e}, \mathbf{i}_V, \boldsymbol{\phi}, \mathbf{a}, \boldsymbol{\pi}) : \mathcal{I} \to \mathbb{R}^{4+1+76744+208499+208499}$.

Simulation of Example 7.5 using BDF of first order with constant step size $h = 6.5 \times 10^{-11}$ on $\mathcal{I} = [0, 1.3 \times 10^{-8}]$, Newton's method as nonlinear solver and hybrid linear solving using `spsolve` and `MUMPS` yields the results in Figure 7.10 for the node potentials. An equivalent circuit of the one in Example 7.5 is for instance given by the one in Figure 7.9

The mock element interface introduced in Section 6.1 also allows for nonlinear materials such as in semiconductor devices of those in Example 7.6. Here the underlying refined model equations are spatially discretized forms of (4.22) or (4.24).

Example 7.6

Consider the CMOS inverter given by the field/circuit coupled problem in Figure 7.12 as from

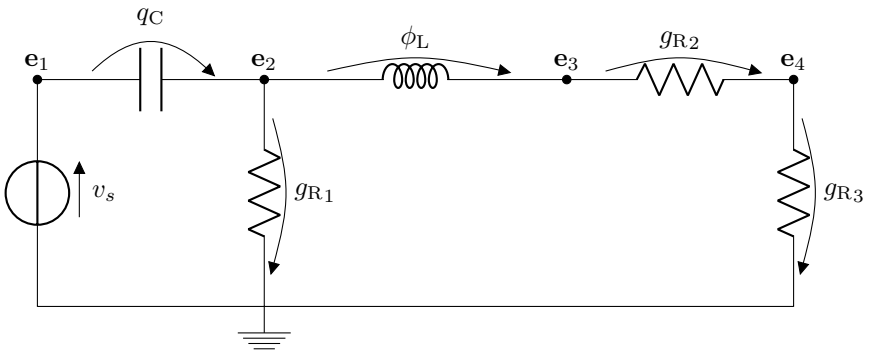


Figure 7.11: Band-pass filter equivalent circuit with $v_s(t) = \sin(\pi 10^8 t) + \sin(2\pi 10^9 t) + \sin(2\pi 0.8 \times 10^{10} t)$, $g_{R1}(t) = g_{R2}(t) = 0.03003$, $g_{R3}(t) = 1$ and $\phi_L(t) = 2.4 \times 10^{-9}$.

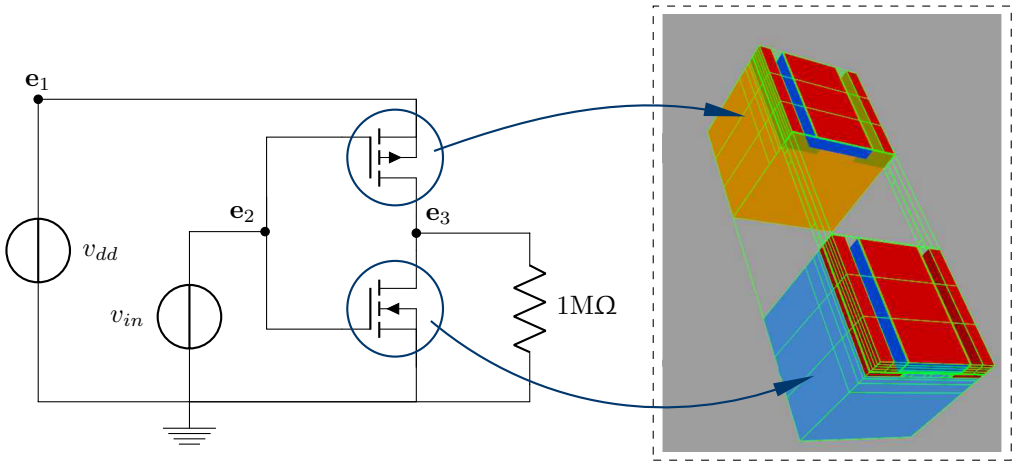


Figure 7.12: CMOS inverter with semiconductor field devices as taken from [Bit+18].

[Bit+18]. Here, the PMOS and NMOS devices are given by 3D EM field models from `DevEM` *and the nonlinear interface is used to couple them into the circuit.*

The nonlinear mock element interface has proven itself to work well with other circuit simulators such as `LinzFrame`[5] from our academic project partner within the *nanoCOPS* project, see [Bit+18; Mat+19].

[5] A circuit simulator of the University of Applied Sciences of Upper Austria.

7.2 Waveform Relaxation Benchmarks

In this section we provide numerical results according to the waveform relaxation method of Gauss-Seidel type introduced in Section 6.3.1. The waveform relaxation is realized in a self-developed flow network DAE framework [Str+18]. This framework, implemented in Python, is adapted to deal with circuit models introduced in Section 3.2 and interfacing also self-developed Python realization of the finite integration technique (FIT) in Section 4.5.
The benchmark examples rely on the coupled field/circuit models presented in Section 2.2 and all satisfy the standard Assumptions 3.7, 3.8, 3.8, 3.8 and 3.10 which already have been used throughout this whole section. Note that the EM field devices replace the mock elements. These assumptions are easy to verify by the choice of benchmark parameters and the circuits' topologies. Moreover, the fulfillment of those assumptions allows for the immediate application of index and convergence theorems already introduced. Note that the convergence plots in the following always show every component whereas their legend is capped due to the restricted space which is indicated by a triple dot.

Systems Settings All the presented example systems are solved monolithically on the one hand and using waveform relaxation with different window sizes on the other hand. Solving the systems in a monolithic fashion is done using the incorporated formulation (5.21). This solution is required in order to obtain a reference solution for the waveform relaxation schemes' error analysis. As for the waveform relaxation schemes, we choose (6.25) and (6.35) as representative variants for the first class of Gauss-Seidel (GS1) and Jacobi type (Jac1) iteration schemes, respectively. As an error measure we take the error of each component which itself is defined by the maximal absolute error over the whole time-interval $\mathcal{I}$.

7.2.1 Low-Pass Filter with Current Source

As for the first benchmark of a coupled circuit and EM device systems, we choose current driven low-pass filters, operating at high frequencies, as given in Figures 7.13 and 7.15. In addition, both systems are index-2 DAEs. In order to be comparable with each other, but in their topology, they share the following setup:

Benchmark Parameters For both circuits, the parameters read $i_s(t) = \sin(2\pi \cdot 10^6 t) + \sin(2\pi \cdot 10^5 t)$ in ampere for the current sources, $q_C = 10^{-6}$ in farad for the capacitor and $\phi_L = 10^{-3}$ in henry for the inductor. Acting as a $100\,\Omega$ resistor, we choose an EM device consisting of a $3 \times 3 \times 802$ micron nickel-chrome (NiCr) bar surrounded by oxide. The material parameters are given in Table 7.1.
The device is embedded into the circuit via contacts at the opposing small facets of the bar. These bounrady parts are attached to respectively $\mathbf{e}_2$ and the circuit's ground. Note that due to the chosen Dirichlet boundary conditions, the complementary boundary can be interpreted as being attached to the circuit's ground node as well.

Solver Settings As for the time integration, during every solving process we use BDF of first order with constant time step size of 10^{-7} seconds on a total time interval $\mathcal{I} = [0, 10^{-5}]$, in seconds. The initial values u_0 and x_0 are set to zero such as the initial guesses $\mathbf{x}^{[0]}$ and, consequently, $\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(\mathbf{x}^{[0]})$ on each discrete time point. Time windowing is not used here, i. e. time integration is realized over the whole time interval.

Example 7.7
Consider the low-pass filter circuit with incorporated EM devices of Figure 7.13.

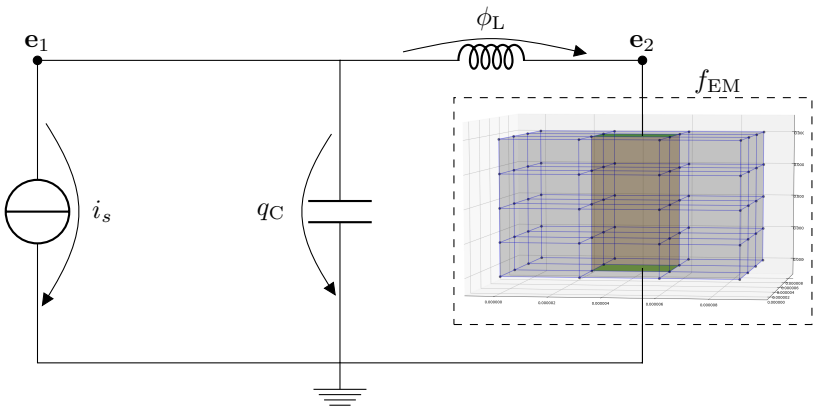


Figure 7.13: Coupled low-pass filter circuit with a current source and an EM device acting as an 100 Ohm resistor.

The quantities' dimensions are $(\mathbf{e}, \mathbf{i}_{\mathrm{L}}, \boldsymbol{\phi}, \mathbf{a}, \boldsymbol{\pi}) : \mathcal{I} \to \mathbb{R}^{2+1+12+52+52}$.

Since all assumptions for Theorem 5.12 are met, the presence of an LE-cutset in the circuit states, that (5.21) of Example 7.7 has index 2.

In fact, the LE-cutset is actually an LE^+-cutset which is why we cannot guarantee convergence of the iteration scheme (6.25) or any other within the same equivalence class GS1 or within Jac1, using Theorem 6.29 and Theorem 6.36, respectively. Indeed, with each iteration, the error of every component magnifies after a head start, see Figure 7.14. In addition, after 7 iterations the step size needed to be adjusted for GS1 in order to overcome difficulties in the linear solving subroutine. After 11 iterations solving was not possible anymore, even with step size adjustments.

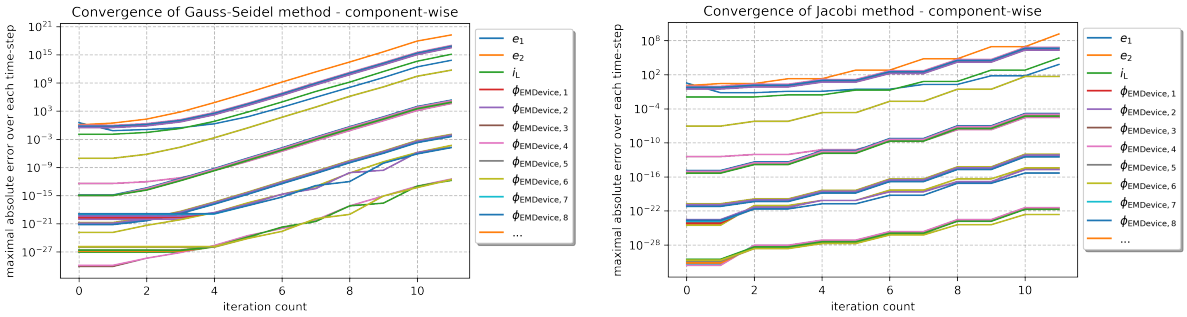


Figure 7.14: Solving the current low-pass filter of Example 7.7 using the Gauss-Seidel iteration scheme (6.25) and Jacobi iteration scheme (6.35) as GS1 and Jac1 representatives. The plot shows the error of each circuit and field variable in terms of iteration count.

Remark 7.8
The stair-like shape of the convergence plots of the Jacobi type WR scheme is due to the fact that the EM subsystem's solution does not change during the first iteration and then an update takes place alternately between the two subsystems while the other remains unchanged.

Example 7.9
The low-pass filter system with incorporated EM devices of Figure 7.15.

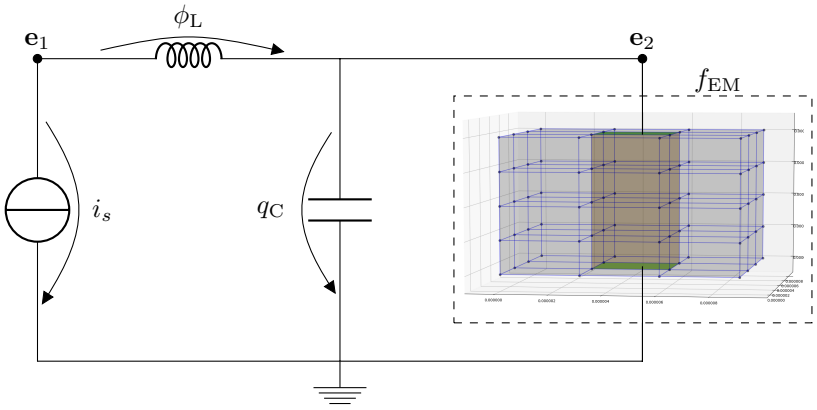


Figure 7.15: Modified coupled low-pass filter circuit with a current source and an EM device acting as an 100 Ohm resistor.

The quantities' dimensions are $(\mathbf{e}, \mathbf{i}_L, \boldsymbol{\phi}, \mathbf{a}, \boldsymbol{\pi}) : \mathcal{I} \to \mathbb{R}^{2+1+12+52+52}$.

Again, all assumptions for Theorem 5.12 are fulfilled. This time it is the presence of an LI-cutset which causes the system (5.21) of Example 7.9 to have at index 2.

Contrary to Example 7.7, we notice the absence of an LIE^+-cutset. As of Corollaries 6.31 and 6.37 , convergence of the Gauss-Seidel type WR scheme (6.25) and Jacobi type WR scheme (6.35), or any other of these equivalence classes, can now be guaranteed. Indeed, Figure 7.16 shows convergence to machine precision after approximately 9 iterations for GS1 and after 18

iterations for Jac1. This is already a quite fast convergence, which can be explained by the relatively high resistance of the EM device of 100 Ohm. If, on the other hand, the conductivity is increased by 10, convergence occurs after twice the number of iterations. Similar observations where made in [Pad20], where a connection of the convergence speed in relation to the resistance along paths between coupling nodes is established but for the circuit subsystem.

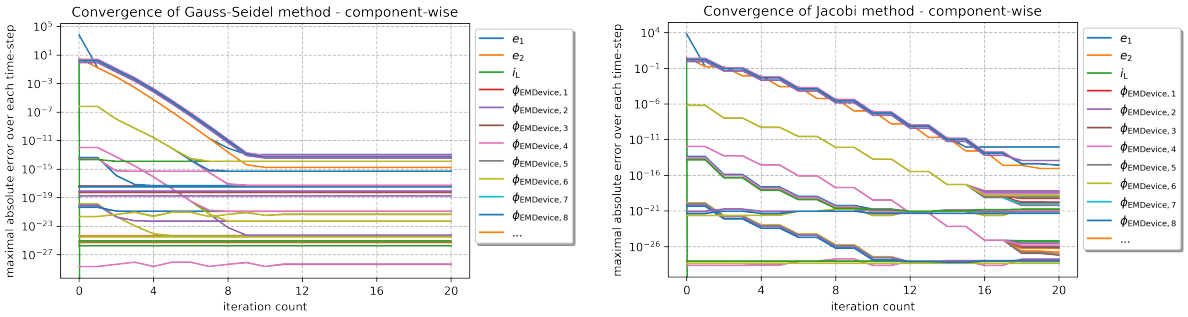


Figure 7.16: Solving the current low-pass filter of Example 7.9 using the Gauss-Seidel iteration scheme (6.25) and Jacobi iteration scheme (6.35) as GS1 and Jac1 representatives. The plot shows the error of each circuit and field variable in terms of iteration count.

Note Remark 7.10 following this example for an explanation concerning the behavior of certain variables close to zero in Figures 7.17 and 7.18.

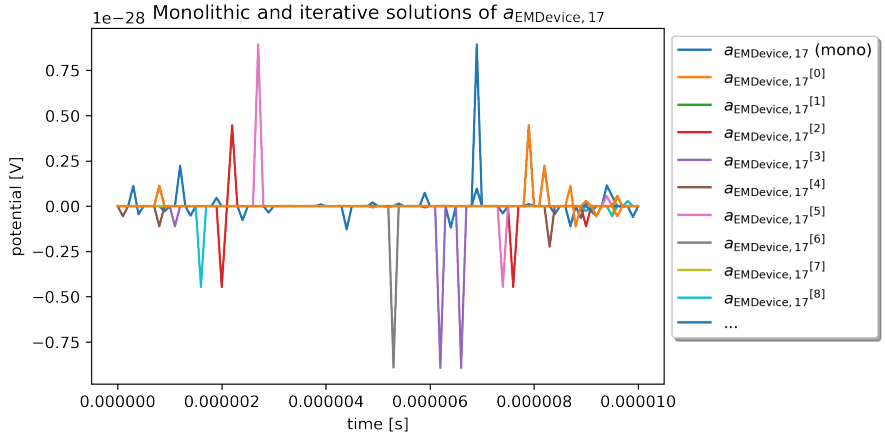


Figure 7.17: Monolithic and iterative solutions of a near zero field variable of the modified low-pass filter in Example 7.9.

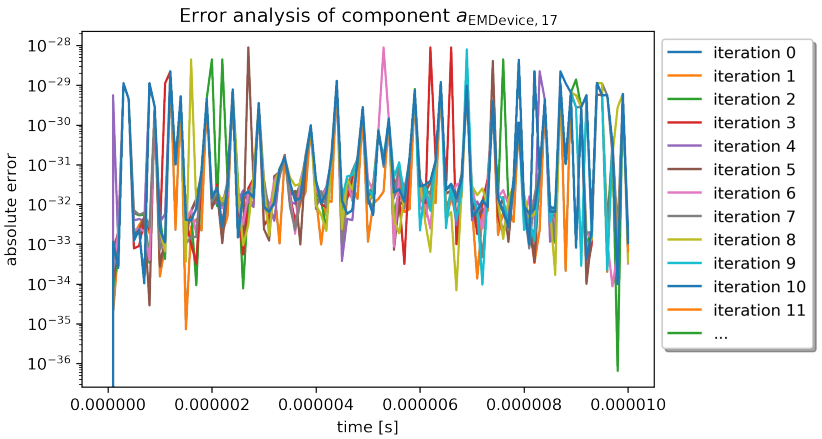


Figure 7.18: Absolute error of iterative solutions of a near zero field variable of the modified low-pass filter in Figure 7.15. Monolithic solution is the reference solution.

Focusing on the filtered signal, given by the node potential $\mathbf{e}_2$, the solutions and errors over the time interval $\mathcal{I}$ can be tracked in terms of iteration count in Figures 7.19 and 7.20.

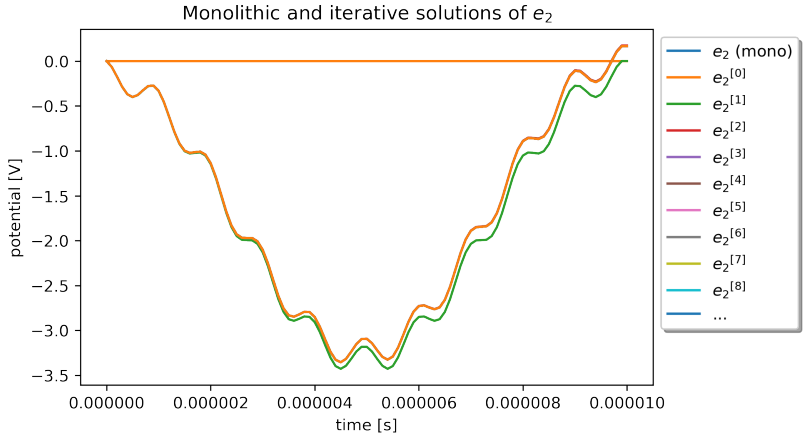


Figure 7.19: Monolithic and iterative solutions of node potential $\mathbf{e}_2$ after the filtering stage of the modified low-pass filter in Example 7.9.

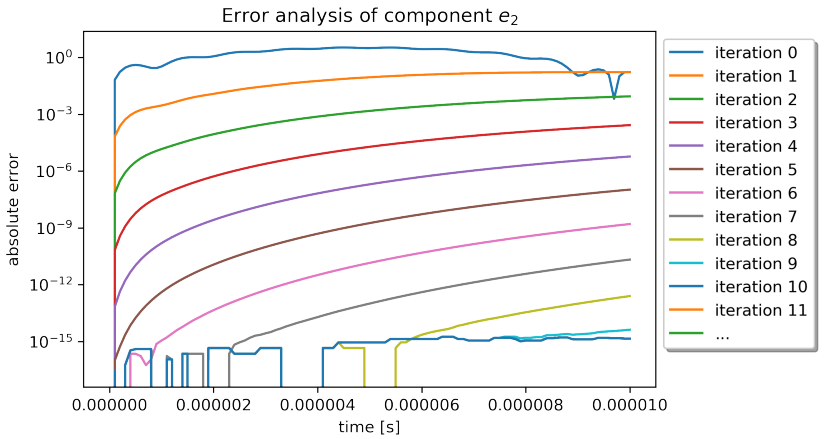


Figure 7.20: Absolute error of iterative solutions of node potential $\mathbf{e}_2$ after the filtering stage of the modified low-pass filter in Example 7.9. Monolithic solution is the reference solution.

Remark 7.10
Some variables are almost zero over time in a first place and contain numerical dirt. Therefore, we do not expect any relative enhancement of the error with increasing iteration count.

7.2.2 Band-Pass Filter using Capacitor and Inductor

As for the last benchmark, we use a capacitor and inductor based band-pass filter structure given in Figure 7.21.

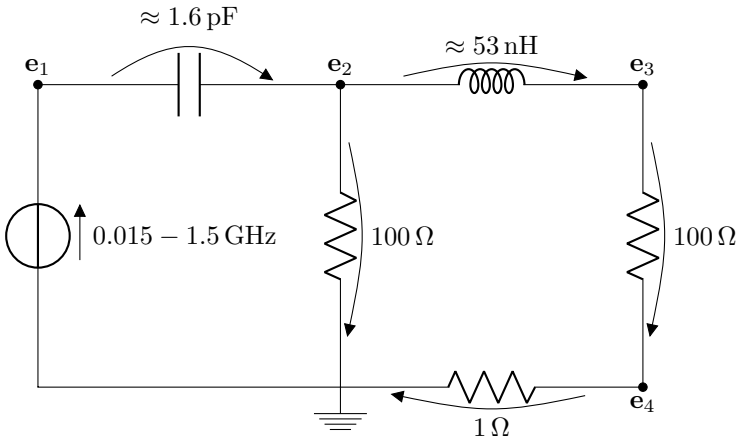


Figure 7.21: Band-pass filter circuit using capacitor and resistor for high- and low-pass filter stages, respectively, and a load of $1\,\Omega$. Pass-band from $0.1 - 0.3\,\mathrm{GHz}$.

This time, the examples are constructed by alternately replacing one of the $100\,\Omega$ resistor with a refined EM device fulfilling the same purpose, see Figures 7.22 and 7.29. Again, the circuit and EM device parts are solved iteratively using Gauss-Seidel type schemes (6.25).

The key message of this benchmark is to emphasize importance of the previous convergence analysis, especially the results collected in Corollary 6.31, when it comes to circuit design decisions involving waveform relaxation methods for solving. In particular, using the topology as a guidance of which elements can safely be replaced by refined elements, in this case the previously introduced full-Maxwell EM field devices, when Gauss-Seidel's method is applied to it.

Benchmark Parameters The parameters used for the circuit read $v_s(t) = \sin(2\pi \cdot 0.015 \cdot 10^9 t) + \sin(2\pi \cdot 0.15 \cdot 10^9 t) + \sin(2\pi \cdot 1.5 \cdot 10^9 t)$ in volt for the voltage sources, $q_\mathrm{C} = 1.5915 \cdot 10^{-11}$ in farad for the capacitor and $\phi_\mathrm{L} = 5.3052 \cdot 10^{-8}$ in henry for the inductor. Acting as a load, we chose a resistor with $g_{\mathrm{R}_\mathrm{load}} = 1$ in ohm. For the filtering resistors we have $g_{\mathrm{R}_1} = 100$, as well in ohm, and the same EM device introduced in 7.2.1, acting as such.
The device is again embedded in the circuit via contacts at the opposing small facets of the bar. These boundary parts are attached to either $\mathbf{e}_2$ and the circuit's ground or $\mathbf{e}_3$ and $\mathbf{e}_4$. Note that due to the chosen Dirichlet boundary conditions, the complementary boundary can be interpreted as being attached to the circuit's ground node as well, explaining why the EM device in Figure 7.29 appears as a three-terminal element.

Solver Settings During every solving process we use BDF of first order with constant time step size of 10^{-11} seconds on a total time interval $\mathcal{I} = [0, 10^{-8}]$, in seconds. The initial values u_0 and x_0 are set to zero such as the initial guesses $\mathbf{x}^{[0]}$ and $\frac{\mathrm{d}}{\mathrm{d}t} d_\mathrm{E}(\mathbf{x}^{[0]})$ on each discrete time point. Time integration is realized over the whole time interval and by making use of time windowing technique.

Example 7.11
Consider the band-pass filter system with incorporated EM devices of Figure 7.22.

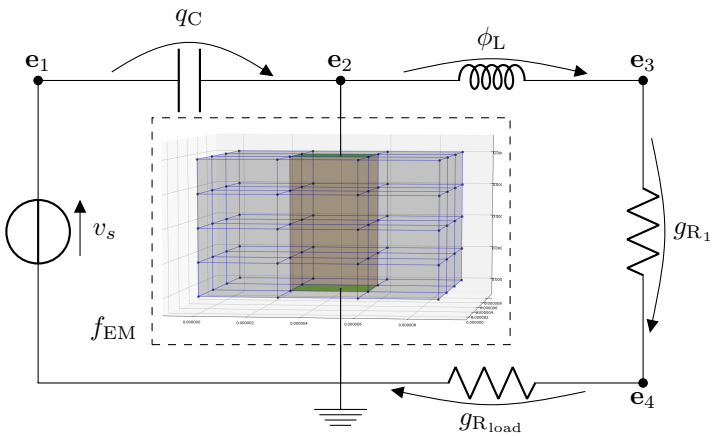


Figure 7.22: Band-pass filter from Figure 7.21 with an EM device acting as an 100 Ohm resistor in the high-pass filter stage.

The quantities' dimensions are $(\mathbf{e}, \mathbf{i}_\mathrm{L}, \mathbf{i}_\mathrm{V}, \boldsymbol{\phi}, \mathbf{a}, \boldsymbol{\pi}) : \mathcal{I} \to \mathbb{R}^{4+1+1+12+52+52}$.

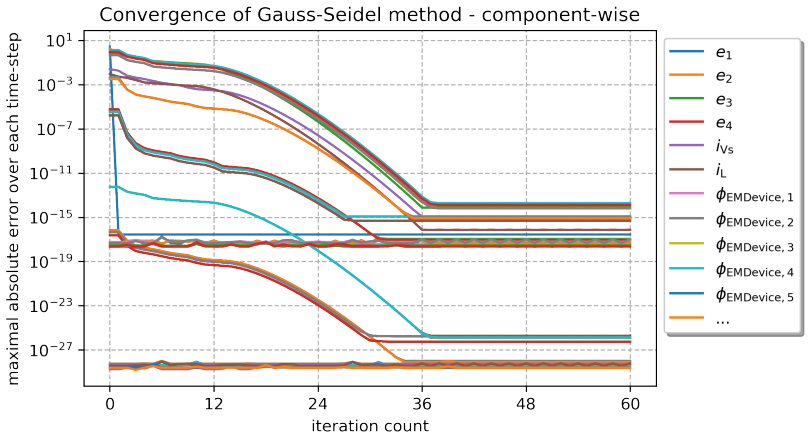


Figure 7.23: Solving the band-pass filter in Example 7.11 using the Gauss-Seidel iteration scheme (6.25). The plot shows the absolute error of each circuit and field variable in terms of Gauss-Seidel iteration count. Convergence after 38 iterations.

Since all assumptions for Theorem 5.12 are met, the absence of CV-loops and LIE-cutsets in the circuit of Example 7.11 states that (5.21) has index 1.

Again, from Corollary 6.31 we expect convergence of the Gauss-Seidel type waveform relaxation method GS1. This is verified by the error plot in Figure 7.23.

The transient simulation results of the node potential $\mathbf{e}_4$, after the filtering stages, can be see in Figure 7.24. We observe faster alignment of the iterative solutions to the monolithic one at the time interval's beginning. This behavior continues to the next high frequencies cycle with the forthcoming iterations and so on. Taking a look at the component's absolute error over the whole time interval of each iteration in Figure 7.25, verifies this observation. Indeed, the error decreases much more accelerated in the very beginning of the time interval $\mathcal{I}$, be aware that this is in a logarithmic scale. Hence, this example seems to be a good candidate for the windowing technique introduced in Section 6.3.

As expected, splitting the time interval $\mathcal{I}$ in 10 equidistant time windows, we already observe convergence for a significant smaller number of iterations. In this case after 16 iterations for GS1, see Figure 7.26. Taking, for example, the joint iterative solutions of node potential $\mathbf{e}_4$, we observe faster convergence on the whole time interval, see Figure 7.27. The improvement becomes clear when looking, for instance, at the very same node potential's error over the time interval in Figure 7.28. We may even ecpect faster convergence by shrinking the time window sizes further.

Remark 7.12
The simple windowing setup in Example 7.11 already saved half the number of time integration steps, which would be necessary without windowing, in order to obtain a solution as good as the monolithic one.

Example 7.13
Consider the band-pass filter system with incorporated EM devices of Figure 7.22.

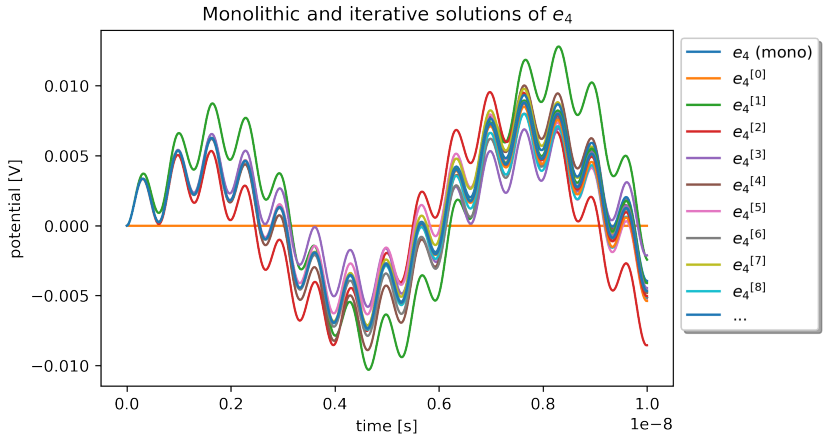


Figure 7.24: Monolithic and iterative solutions of node potential $\mathbf{e}_4$ after the filtering stages of the band-pass filter's first variant in Figure 7.22.

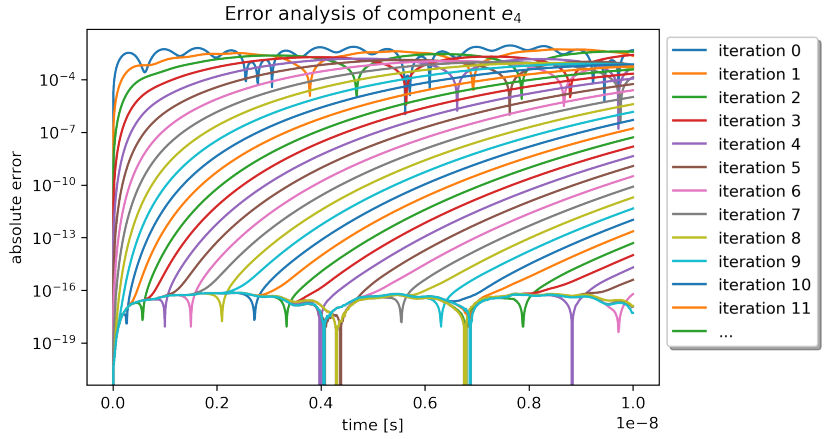


Figure 7.25: Absolute error of iterative solutions of node potential $\mathbf{e}_4$ after the filtering stages of the band-pass filter's first variant in Figure 7.22. Monolithic solution is the reference solution.

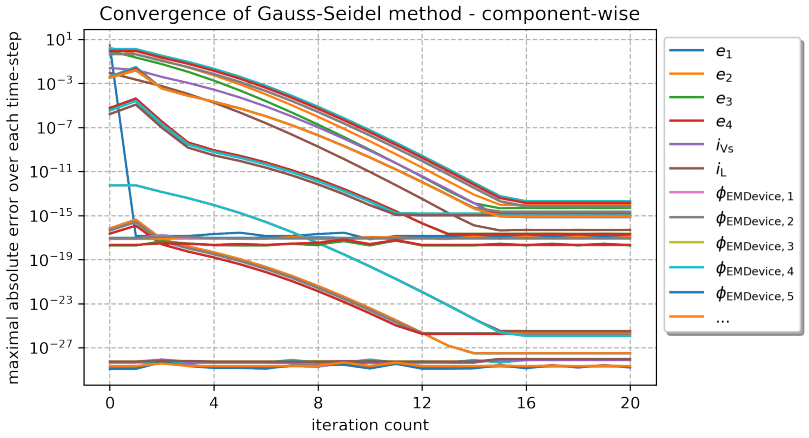


Figure 7.26: Solving the band-pass filter in Figure 7.22 using the Gauss-Seidel iteration scheme (6.25) and windowing. The plot shows the absolute error of each circuit and field variable in terms of iteration count. Convergence after 16 iterations.

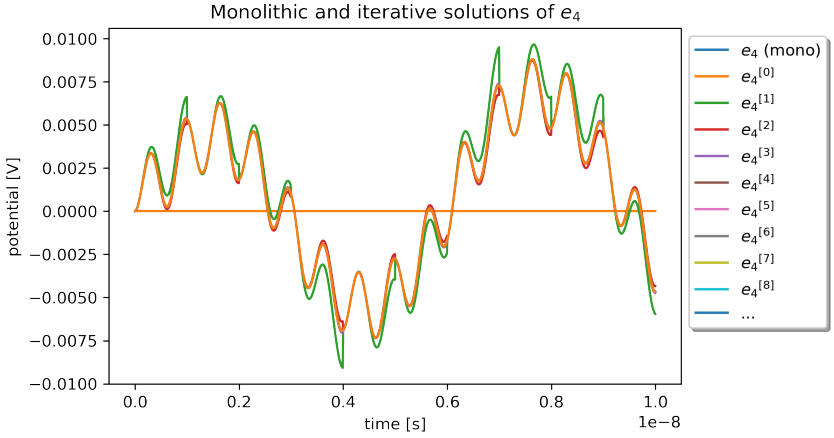


Figure 7.27: Monolithic and iterative solutions of node potential $\mathbf{e}_4$ after the filtering stages of the band-pass filter's first variant in Figure 7.22 using windowing.

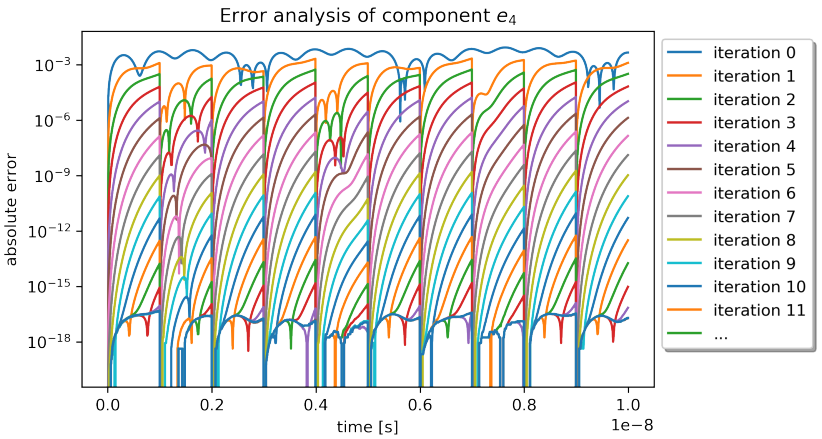


Figure 7.28: Absolute error of iterative solutions of node potential $\mathbf{e}_4$ after the filtering stages of the band-pass filter's first variant in Figure 7.22. Monolithic solution is the reference solution and windowing was used.

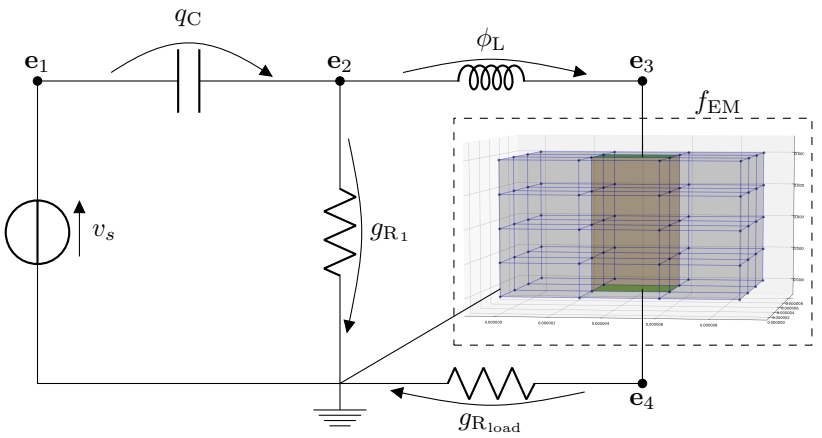


Figure 7.29: Band-pass filter from Figure 7.21 with an EM device acting as an 100 Ohm resistor in the low-pass filter stage.

The quantities' dimensions are $(\mathbf{e}, \mathbf{i}_L, \mathbf{i}_V, \boldsymbol{\phi}, \mathbf{a}, \boldsymbol{\pi}) : \mathcal{I} \to \mathbb{R}^{4+1+1+12+52+52}$.

Contrary to Example 7.11, there is now an LE-cutset in the circuit of Example 7.13 stating, that (5.21) has index 2, by Theorem 5.12.

Whereas the index-2 property is not necessary an obstacle to convergence of (6.25), see Examples 7.7 and 7.9, it is the very fact that the EM device is part of this LE-cutset and, therefore, the presence of an LE^+-cutset. In other words, convergence cannot be guaranteed using Corollary 6.31. Indeed, the Gauss-Seidel method diverges, as of Figure 7.30. Note, after 7 iterations the step size needed to be adjusted in order to overcome difficulties in the linear solving subroutine which, nevertheless, forced the time integration to shut down after 11 iterations.

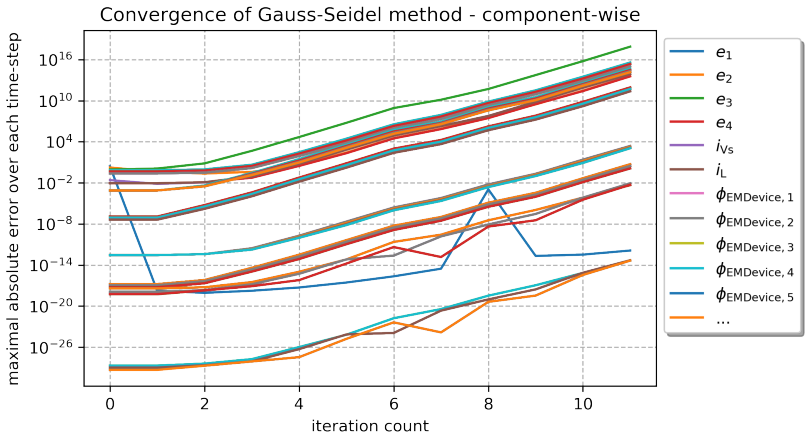


Figure 7.30: Solving the band-pass filter in Figure 7.29 using the Gauss-Seidel iteration scheme (6.25). The plot shows the absolute error of each circuit and field variable in terms of Gauss-Seidel iteration count. Divergent till shutdown after 11 iterations.

Having a look at a single component, e. g. $\mathbf{e}_4$ in Figure 7.31, we do not expect much gain from using windowing for this example.

7.3 Conclusions

In this chapter we focused on the simulation of some concrete field/circuit couped problems. In the first section we dealt with monolithic simulation approach. For a simple toy Example 7.1, using self-implemented FIT and circuit simulator, we have shown that a monolithic simulation is possible. Based on this example we have verified, that certain techniques such as scaling and structural exploitation, introduced in Chapter 6, can have a positive effect on the linear solving routine when using iterative methods. For the same EM device, simulations using the model equations introduced in this treatise and those obtained by `DevEM` yield comparable results, verifying the successful implementation of the interface. Furthermore, it was finally possible to simulate higher-dimensional models, such as the ACCO antenna or inductor prototypes in Table 7.5, in a circuit environment by making use of the techniques elaborated in Section 6.2, e. g. the $\mathbf{a}/\boldsymbol{\pi}$-elimination and hybrid linear solver approach. Also circuits with multiple EM device and semiconductor device models could be simulated, see for instance the band-pass filter in Example 7.5 or the CMOS inverter in Example 7.6.
The second part of this chapter dealt with the waveform relaxation approach to field/circuit problems. In order to support the theorems derived in Chapter 6, we used refined model equations from the self-implemented FIT. We demonstrated, that WR schemes can diverge if the topological criteria which guarantee convergence are violated, for instance the presence of an LIE^+-cutset when using GS1 according to Theorem 6.29. Vice versa, if they are met, we observed convergence. Further, we verified that using the windowing technique, as introduced in Section 6.3, can improve convergence by magnitudes.

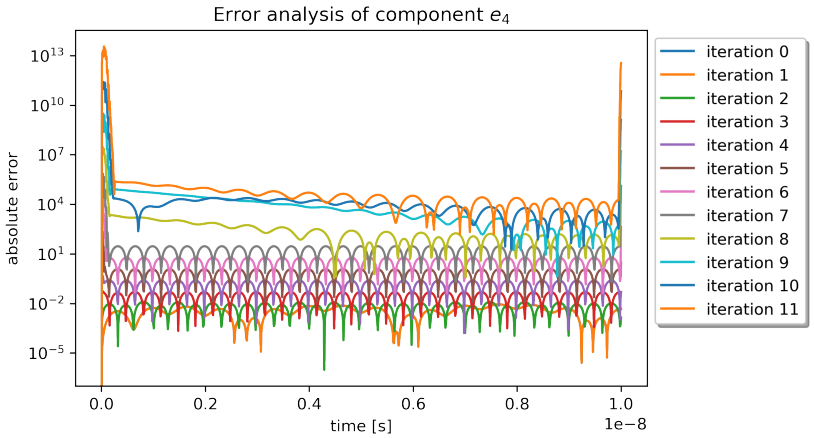


Figure 7.31: Absolute error of iterative solutions of node potential $\mathbf{e}_4$ after the filtering stages of the band-pass filter's first variant in Figure 7.29. Monolithic solution is the reference solution.

8 Summary and Outlook

Summary In this treatise we considered modeling aspects and numerical analysis of circuits coupled to electromagnetic (EM) field devices with respect to their simulation. The conceptional basis was laid down by the introduction of a mock element in Chapter 3, dealing as an interface between circuitry with classical lumped element models, as they are used by the modified nodal analysis (MNA), and refined models, arising by spatially discretized partial differential equation systems. In particular with regard to EM field devices modeled using full-wave Maxwell's equations in Lorenz-gauged $\mathbf{A}-\boldsymbol{\varphi}$ formulation, for which the finite integration technique is used as a spatially discretization method, see Chapter 4. Introducing a mock element interface was motivated by the collaboration with the company *MAGWEL* as we intended to couple their device-electro-magnetic modeler `DevEM` to circuit simulators based upon MNA. With the help of this mock element interface, generic field/circuit couplings could be established in different ways favoring the MNA philosophy, that is incorporating EM field devices as current controlling lumped elements with additional intrinsic equations, see Chapter 5. To be more precise, we explored several versions of coupling the EM device model into the circuit model using the introduced mock element interface. They include forms that interpret parts of the discretized Maxwell's field equations as equivalent circuit.

The numerical analysis of these coupled systems' differential-algebraic equations (DAEs) has shown that their index does not exceed 2. Though considering different EM field device models, correlations with the circuit's topology comply with the ones, developed in [Bau12; SBD12; Jan15; Cor+20].

Concerning the simulation of coupled systems, see Chapter 6, we first provided a framework of how to realize the coupled modeling with regard to model equations that are only accessible via function evaluations. Hence, support for third party black-box tools for EM field devices can be established enabling the field/circuit coupled simulation. We discussed time integration, scaling methods, structural exploitation and a hybrid approach of solving underlying linear systems of equations which enabled the use of specialized solvers for each subsystem. All this together resulted in a successful simulation of coupled field/circuit systems, also with `DevEM` test cases we considered during our work with *nanoCOPS* project partners, see Chapter 7. Albeit the `DevEM` models are black-box models, all necessary information is exchanged by the interface to apply these methods.

Further, we studied a monolithic simulation approach and waveform relaxation (WR) methods in a comparing manner. In contrast to the monolithic approach, various coupled system formulations differ when WR methods are applied to them, as shown in Section 6.3. Concerning the WR approach, a convergence analysis of all considered coupling forms was offered which extends the ones in [AG01; Sch+10a; PT18] to full-Maxwell EM field devices with a more general

coupling structure that involves additional derivatives. This analysis includes a convergence theorem allowing up to two steps of pre-iterated solutions, cf. [PT18; JK96; Sau19]. For all the field/circuit coupled variants, criteria were developed which guarantee convergence for both Gauss-Seidel and Jacobi type WR schemes.
For the here considered model equations, the criteria are topology dependent only and based on the theorems developed in Chapter 3. This is of advantage when simulating, for example, a new design for an EM device in its specific circuitry environment of application. If WR methods are used, engineers can take into account the topological position of the device in their design decisions and justify the choice of device models to ensure a successful simulation with WR methods. Otherwise, a monolithic simulation approach is recommended. Moreover, given access to specialized linear solvers for the EM subsystems, the monolithic simulation profits from a hybrid solving approach, similar to WR methods on a system level but for the underlying linear equation systems. Furthermore, we have shown that due to the coupling structure this hybrid method is bounded in complexity. Further, regardless of whether a monolithic or co-simulation approach was chosen, there exist structural exploits advancing the simulation for all variants such as the $\mathbf{a}/\boldsymbol{\pi}$-elimination or scaling methods see Chapter 6.

This treatise finishes with some numerical benchmarks in Chapter 7 underlining the simulation methods and convergence results. For the latter ones, precise model equations were required which necessitated the self-implementation of an EM modeler, realized in `python`. Therefore, the benchmarks are produced by a self-implemented electric circuit simulator interfacing both, the self-implemented version of an EM device modeler and `DevEM` using the same mock element based interface design. We verified the monolithic simulation's success for test cases using the here presented model equations and those obtained by a black-box modeler. In the second part, we performed benchmarks supporting the topological criteria for WR methods that demonstrate convergence where it is to be expected and divergence where criteria are violated.

Outlook The introduced mock element interface is not restricted to EM device models. It also allows the integration of other, possibly black-box DAE, models. In this way the analysis, especially with regards to convergence of WR methods, can be extended to a mock element description allowing also for DAEs.
To take this idea even to a further level, one can think of convergence analysis for generalized circuit elements, see [Cor+20], where classical circuit elements and certain refined models are grouped into categories according to their impact on numerical and analytical difficulties regarding the overall system's index. In fact, the here presented refined model for the EM field device using the full-Maxwell approach influenced the design of generalized element descriptions. Not only the EM device, but the whole mock element fits into this concept as an inductance-like element. Accordingly, the topological index statements in [CDS19] are confirmed by Theorem 5.12 and Corollary 5.14.
As a next step we propose to do this convergence analysis for an inductance-like element, separated as an individual subsystem from the circuit, first and continue with capacitance-like and resistance-like elements. In this way, convergence criteria can be provided which not only apply to refined models based on Maxwell's equations, alternative formulations or approximations thereof. In concrete, this seems also to be a promising approach to propose design decisions for

the coupling of further complex models such as gas networks, water networks, power networks etc. when they are intended to be solved using WR methods. For instance such gas/circuit coupled problems which are part of the power-to-gas modeling and solving methodology considered by the *MathEnergy*[1] project. Consider the following example of a gas network coupled to a circuit system from [Cle+20].

Example 8.1
Given the simple power circuit coupled to gas network via an electrolyser in Figure 8.1, the gas

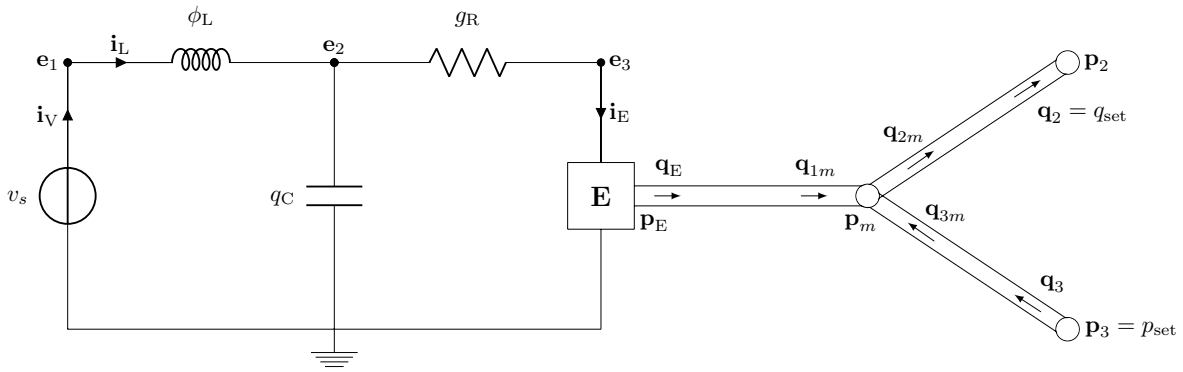


Figure 8.1: Example of power circuit to gas network coupling via an electrolyser.

network benchmark DAE reads

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{p}_{\mathrm{E}} + a_1(\mathbf{q}_{1m} - \mathbf{q}_{\mathrm{E}}) &= 0,\\
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{p}_2 + a_2(\mathbf{q}_{2m} - \mathbf{q}_2) &= 0,\\
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{p}_m + a_3(\mathbf{q}_3 - \mathbf{q}_{3m}) &= 0,\\
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q}_{1m} + b_1(\mathbf{p}_m - \mathbf{p}_{\mathrm{E}}) &= -g(\mathbf{p}_{\mathrm{E}}, \mathbf{q}_{1m}),\\
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q}_{2m} + b_2(\mathbf{p}_m - \mathbf{p}_2) &= -g(\mathbf{p}_2, \mathbf{q}_{1m}),\\
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q}_3 + b_3(\mathbf{p}_3 - \mathbf{p}_m) &= -g(\mathbf{p}_m, \mathbf{q}_3),\\
\mathbf{p}_3 &= p_{\mathrm{set}}(t),\\
\mathbf{q}_2 &= q_{\mathrm{set}}(t),\\
\mathbf{q}_{1m} - \mathbf{q}_{2m} + \mathbf{q}_{3m} &= 0,
\end{aligned}
$$

for given parameters $a_1, a_2, a_3, b_1, b_2, b_3 \in \mathbb{R}$ *and functions* $p_{\mathrm{set}}, q_{\mathrm{set}} : \mathbb{R} \to \mathbb{R}$ *and* $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. *Note that the parameters are short-handedly written and actually involve diameters, cross-sectional areas, lengths of the pipes, friction and more. The DAE reflects spacially discretized*

[1] Federal Ministry for Economic Affairs and Energy (BMWi) supported project *MathEnergy* under grant 0324019E

pipe equations for gas transport using a simple topology adopted spatial discretization, see [Huc18]. The coupling equation is given by the chemical reaction

$$\mathbf{q}_{\mathrm{E}} = \gamma \mathbf{i}_{\mathrm{E}}$$

with $\gamma \in \mathbb{R}$ *The power circuit DAE following the MNA approach reads*

$$\begin{aligned} \mathbf{i}_{\mathrm{L}} - \mathbf{i}_{\mathrm{V}} &= 0, \\ C\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{e}_2 - \mathbf{i}_{\mathrm{L}} + G(\mathbf{e}_2 - \mathbf{e}_3) &= 0, \\ G(\mathbf{e}_3 - \mathbf{e}_2) &= -\mathbf{i}_{\mathrm{E}}, \\ L\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{i}_{\mathrm{L}} &= \mathbf{e}_1 - \mathbf{e}_2, \\ v_s(t) + \mathbf{e}_1 &= 0, \end{aligned}$$

for given $C, L \in \mathbb{R}$ *and* $G, v_s : \mathbb{R} \to \mathbb{R}$. *The coupling equation arises from power balance* $\mu \mathbf{i}_{\mathrm{E}} \cdot u_{\mathrm{E}} = \frac{1}{\varrho}(\mathbf{p}_{\mathrm{E}} - p_o) * \mathbf{q}_{\mathrm{E}}$ *with* $u_{\mathrm{E}} = \mathbf{e}_3$ *regarding a priori known* μ, ρ *and* $p_o \in \mathbb{R}$. *Due to* $\mathbf{q}_{\mathrm{E}} = \gamma \mathbf{i}_{\mathrm{E}}$ *we obtain*

$$\mu \mathbf{e}_3 = \frac{\gamma}{\varrho}(\mathbf{p}_{\mathrm{E}} - p_o).$$

Note that the incidence matrices read

$$A_{\mathrm{C}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad A_{\mathrm{R}} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad A_{\mathrm{V}} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \quad A_{\mathrm{L}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad A_{\mathrm{E}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Incorporation of the source and coupling terms, the gas network DAE becomes an ODE that is, for $\mathbf{u} = (\mathbf{p}_{\mathrm{E}}, \mathbf{p}_2, \mathbf{p}_m, \mathbf{q}_{1m}, \mathbf{q}_{2m}, \mathbf{q}_3)$, $\mathbf{x} = (\mathbf{i}_{\mathrm{L}}, \mathbf{i}_{\mathrm{V}}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, $\mathbf{e} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ *and* $\mathbf{w} = (\mathbf{w}_{\mathrm{C}}, \mathbf{w}_{\mathrm{L}})$,

$$M_{\mathrm{E}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} + b_{\mathrm{E}}(\mathbf{u}, t) = c_{\mathrm{E}}(A_{\mathrm{E}}^{\top}\mathbf{e}), \tag{8.1}$$

with

$$M_{\mathrm{E}} = I, \quad b_{\mathrm{E}}(\mathbf{u}, t) = \begin{pmatrix} a_1 \mathbf{q}_{1m} \\ a_2(\mathbf{q}_{2m} - q_{set}(t)) \\ a_3(\mathbf{q}_3 + \mathbf{q}_{1m} + \mathbf{q}_{2m}) \\ b_1(\mathbf{p}_m - \mathbf{p}_{\mathrm{E}}) + g(\mathbf{p}_{\mathrm{E}}, \mathbf{q}_{1m}) \\ b_2(\mathbf{p}_m - \mathbf{p}_2) + g(\mathbf{p}_2, \mathbf{q}_{1m}) \\ b_3(p_{\mathrm{set}}(t) - \mathbf{p}_m) + g(\mathbf{p}_m, \mathbf{q}_3) \end{pmatrix}, \quad c_{\mathrm{E}}(\mathbf{i}_{\mathrm{E}}) = \begin{pmatrix} a_1 \gamma \mathbf{i}_{\mathrm{E}} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Then, the gas network can be interpreted as a voltage controlling pendant to the mock element with (8.1) *being the intrinsic equation, substituting* (3.6), *and the coupling equation*

$$A_{\mathrm{E}}^{\top}\mathbf{e} = f_{\mathrm{E}}(\mathbf{u}) := \frac{\gamma}{\varrho\mu}(\mathbf{p}_{\mathrm{E}} - p_o)$$

substituting the constitutive equation (3.5) *in a voltage controlling manner.*

This example shows one way in which the modeling aspects and numerical analysis of this treatise can be extended to other forms of coupling. In particular, by considering further voltage controlling mock element interfaces in addition to the current controlling one and finally to all generalized elements. Furthermore, Theorem 6.19 might turn out to be useful in developing new schemes beyond Jacobi and Gauss-Seidel type or generalizing damping or vector iterations as subject to further investigation.

A Appendix

A.1 Electromagnetism

Lemma A.1
Let $\Omega \subset \mathbb{R}^3$ be a simply connected spatial domain, $\Gamma_E \subset \partial\Omega$ open part of the boundary and $\mathcal{I} \subset \mathbb{R}$ a time interval. Assuming electric boundary conditions (4.13) on Γ_E and the MEs (4.1) to hold everywhere, we obtain

$$\mathbf{B} \cdot n = 0, \quad \text{on } \Gamma_E \times \mathcal{I}$$

if once fulfilled at some point $t \in \mathcal{I}$.

Proof. The boundary conditions (4.13) state that $\mathbf{E} \times n = 0$ on $\Gamma \times \mathcal{I}$. Further we have from the Maxwell's equations (4.1) that $-\frac{\partial}{\partial t}\mathbf{B} = \nabla \times \mathbf{E}$. Scalar multiplication with the outer unit normal n from the left side yields:

$$-n \cdot \frac{\partial}{\partial t}\mathbf{B} = n \cdot (\nabla \times \mathbf{E}) \quad \text{on } \Gamma \times \mathcal{I}$$

From vector calculus we have that

$$-n \cdot \frac{\partial}{\partial t}\mathbf{B} = \nabla \cdot (\underbrace{\mathbf{E} \times n}_{=0 \text{ on } \Gamma}) = 0, \quad \text{on } \Gamma_E \times \mathcal{I}$$

Since one $t \in \mathcal{I}$ satisfies $\mathbf{B} \cdot n = 0$ on Γ_E, due to the assumption, this relation also holds on whole $\mathcal{I}$. □

A.2 Graph Theory

In the following we provide common definitions and lemmata concerning graph theory for the sake of electric circuit modeling and analysis. Most of them are taken from [Tis04] and [ST20] or are adoptions to them regarding basis functions instead of projectors.

Definition A.2 (loop)
A subgraph $\mathcal{G}_l$ of a connected graph $\mathcal{G}$ is called a loop if $\mathcal{G}_l$ is connected and each node of $\mathcal{G}_l$ connects exactly two branches of $\mathcal{G}_l$.
We say $\mathcal{G}$ has an X-loop if there exists a loop whose branches are of type X.

Definition A.3 (tree)
A subgraph $\mathcal{G}_t$ of a connected graph $\mathcal{G}$ is called a tree if $\mathcal{G}_t$ is connected, contains all nodes of $\mathcal{G}$ and has no loops.
We say $\mathcal{G}$ has an X-tree if there exists a tree whose branches are of type X.

Definition A.4 (cutset)
A subset $\mathcal{B}_c \subset \mathcal{B}$ of a connected graph $\mathcal{G} = (\mathcal{N}, \mathcal{B})$ is called a cutset if $(\mathcal{N}, \mathcal{B}\backslash\mathcal{B}_c)$ is an unconnected graph and $(\mathcal{N}, \mathcal{B}\backslash\mathcal{B}_c + \{b\})$ for $b \in \mathcal{B}_c$ is again a connected graph.
We say $\mathcal{G}$ has an X-cutset if it there exists a cutset whose branches are of type X.

Lemma A.5 (tree)
Let A be the (reduced) incidence matrix of a connected graph $\mathcal{G} = (\mathcal{N}, \mathcal{B})$ and with $n = |\mathcal{N}|$ nodes. Then, $n-1$ columns of A are linear independent if and only if the branches of these columns form a tree.

Proof. See e. g. Theorem A.3 in [Tis04]. □

Lemma A.6 (loops and cutsets)
Let $\mathcal{G}$ be a connected graph with node set $\mathcal{N}$ and branch set $\mathcal{B}$. Assume that there is a non-empty $\mathcal{B}_\mathrm{X} \subsetneq \mathcal{B}$, e. g, found by collecting all branches of a certain types X. With A_X and A_Y we denote the (reduced) incidence matrices of the subgraphs obtained by the same note set $\mathcal{N}$ and branch sets $\mathcal{B}_\mathrm{X}$ and $\mathcal{B}\backslash\mathcal{B}_\mathrm{X}$, respectively. Then it holds

1. *$\mathcal{G}$ has an X-loop if and only if $\ker A_\mathrm{X} \neq \{0\}$.*
2. *$\mathcal{G}$ has an X-cutset if and only if $\ker A_\mathrm{Y}^\top \neq \{0\}$.*

Proof. To 1.: The proof is given by [Tis04, Thm. A.2].
To 2.: The proof is basically similar to [Tis04, Lemma 1.2].
Let n be the number of nodes in $\mathcal{N}$ and, without loss of generality, let the reference node have the highest number.

"⇒": From $\mathcal{G}$ having an X-cutset follows, that $\mathcal{G}_Y = (\mathcal{N}, \mathcal{B}\backslash\mathcal{B}_\mathrm{X})$ is unconnected. That means that there are at least two disjoint node sets that are not connected. We sort the nodes in $\mathcal{N}$ such that the first $n_0 (> 1)$ nodes do not belong to the connected component containing the reference node. Similar, we sort the branches in $\mathcal{B}\backslash\mathcal{B}_\mathrm{X}$ yielding the incidence matrix

$$A_\mathrm{Y} = \begin{bmatrix} * & 0 \\ 0 & * \end{bmatrix} \begin{matrix} n_0 \\ n-1-n_0 \end{matrix}.$$

The upper left block of A_Y is representing the full incidence matrix of a graph and has at most row rank $n_0 - 1$, depending on how many connected components there are, as this scheme can be successively applied. The lower right block is representing the (reduced) incidence matrix of a connected graph and therefore has row rank $n-1-n_0$. In total A_Y has at most row rank $n-2$ with $n-1$ rows. Hence, A_Y has no full row rank and thus $\ker A_\mathrm{Y}^\top \neq \{0\}$.

"⇐": From $\ker A_{\mathrm{Y}}^{\top} \neq \{0\}$ we directly follow that A_{Y} has no full row rank. As $\mathcal{G}$ is connected by assumption, there must be a minimal subset of branches $\mathcal{B}_c \subset \mathcal{B}_{\mathrm{X}}$ such that

$$\begin{bmatrix} A_{\mathcal{B}_c} & A_{\mathrm{Y}} \end{bmatrix}$$

has row rank $n-1$. Removing each $b \in \mathcal{B}_c$ from $\mathcal{B} \backslash \mathcal{B}_{\mathrm{X}} \cup \mathcal{B}_c$ would result in a lower row rank. Hence, with $\mathcal{B}_c$ we found an X-cutset. □

Lemma A.7
Let $\mathcal{G}$ be a connected graph with node set $\mathcal{N}$ and branch set $\mathcal{B}$. Assume $\mathcal{B}_{\mathrm{X}} \subset \mathcal{B}$ and $\mathcal{B}_{\mathrm{Y}} \subset \mathcal{B}$ to be branch subsets for some branch types X and Y, respectively. With A_{X} and A_{Y} we denote the (reduced) incidence matrices of the subgraphs $(\mathcal{N}, \mathcal{B}_{\mathrm{X}})$ and $(\mathcal{N}, \mathcal{B}_{\mathrm{Y}})$, respectively. With this given, let Q_{X} be a basis (function) of the kernel of A_{X}, i. e. $\operatorname{im} Q_{\mathrm{X}} = \ker A_{\mathrm{X}}^{\top}$.
Then, the matrix (function) $Q_{\mathrm{X}}^{\top} A_{\mathrm{Y}}$ has full column rank if and only if there is no XY-loop containing at least one Y-type branch.

Proof. The proof is basically similar to [Tis04, Lemma 1.3].
"⇒": Assume $\mathcal{G}$ has an XY-loop. By Lemma A.6 it follows that $\ker \begin{bmatrix} A_{\mathrm{X}} & A_{\mathrm{Y}} \end{bmatrix} \neq \{0\}$. Therefore, it exists a non-zero (x, y) such that

$$A_{\mathrm{X}} x + A_{\mathrm{Y}} y = 0 \,.$$

Left multiplying the latter expression by $Q_{\mathrm{X}}^{\top}$ lets the first summand vanish and consequently we obtain $Q_{\mathrm{X}}^{\top} A_{\mathrm{Y}} y = 0$. As $y \neq 0$ it follows, that $Q_{\mathrm{X}}^{\top} A_{\mathrm{Y}}$ has no full columns rank.

"⇐": Assume that $Q_{\mathrm{X}}^{\top} A_{\mathrm{Y}}$ has no full columns rank. Then, there exists a $y \neq 0$ such that $Q_{\mathrm{X}}^{\top} A_{\mathrm{Y}} y = 0$. Thus, $A_{\mathrm{Y}} y \in \ker Q_{\mathrm{X}}^{\top}$. As, by definition, $\ker Q_{\mathrm{X}} = \operatorname{im} A_{\mathrm{X}}$, there exists an x such that

$$A_{\mathrm{X}} x + A_{\mathrm{Y}} y = 0 \,.$$

Even if x might be zero, e. g. when A_{Y} itself has no full column rank and y was chosen s. t. $A_{\mathrm{X}} y = 0$, it still holds $y \neq 0$ and therefore there exists and XY-loop with at least one Y-type branch. □

Lemma A.8
Let $\mathcal{G}$ be a graph with node set $\mathcal{N}$ and branch set $\mathcal{B}$. Assume $\mathcal{B}_{\mathrm{X}i} \subset \mathcal{B}$, for $i = 1, \ldots, n$, to be branch subsets for $n \in \mathbb{N}$ branch types X_i, respectively, and type Z refers to the complementary branch subset $\mathcal{B}_{\mathrm{Z}} = \mathcal{B} \backslash \left(\bigcup_{i=1}^{n} \mathcal{B}_{\mathrm{X}i}\right)$. With $A_{\mathrm{X}i}$ we denote the (reduced) incidence matrices of the subgraphs $(\mathcal{N}, \mathcal{B}_{\mathrm{X}i})$, respectively. With this given, let $Q_{\mathrm{X}1}$ be a basis (function) of the kernel of $A_{\mathrm{X}1}^{\top}$, i. e. $\operatorname{im} Q_{\mathrm{X}1} = \ker A_{\mathrm{X}1}^{\top}$. Successively, let $Q_{\mathrm{X}k}$, for $k = 2, \ldots, n-1$ be a basis (function) of the kernel of $A_{\mathrm{X}k}^{\top} \prod_{i=1}^{k-1} Q_{\mathrm{X}i}$.
Then, the matrix (function) $Q_{\mathrm{X}n-1}^{\top} \ldots Q_{\mathrm{X}1}^{\top} A_{\mathrm{X}n}$ has full row rank if and only if there is no Z-cutset.

Proof. "⇒": Assume $\mathcal{G}$ has a Z-cutset. From Lemma A.6 we deduce that there exists a non-zero y such that

$$A_{\mathrm{X}i}^{\top} y = 0 \,, \text{ for all } i = 1, \ldots, n \,.$$

As $y \in \ker A_{\mathrm{X}1}^\top$, it has, by definition, to be in the image of $Q_{\mathrm{X}1}$ as well whereby $Q_{\mathrm{X}1}$ has full column rank. Therefore, it exists a non-zero y_1 such that $y = Q_{\mathrm{X}1} y_1$. That gives us

$$0 = A_{\mathrm{X}2}^\top y = A_{\mathrm{X}2}^\top Q_{\mathrm{X}1} y_1$$

implying that $y_1 \in \ker A_{\mathrm{X}2} Q_{\mathrm{X}1}^\top = \operatorname{im} Q_{\mathrm{X}2}$. Again, since $Q_{\mathrm{X}2}$ has full column rank, it exists a non-zero y_2 such that $y_1 = Q_{\mathrm{X}2} y_2$ yielding

$$0 = A_{\mathrm{X}3}^\top y = A_{\mathrm{X}3}^\top Q_{\mathrm{X}1} Q_{\mathrm{X}2} y_2 \ .$$

With the same argumentation, we successively derive a non-zero y_{n-1} such that

$$0 = A_{\mathrm{X}n} \prod_{i=1}^{k-1} Q_{\mathrm{X}i} \cdot y_{n-1} \ .$$

Hence, the matrix (function) $Q_{\mathrm{X}n-1}^\top \dots Q_{\mathrm{X}1}^\top A_{\mathrm{X}n}$ cannot have full row rank.

"⇐": Assume that $Q_{\mathrm{X}n-1}^\top \dots Q_{\mathrm{X}1}^\top A_{\mathrm{X}n}$ has full row rank. Then, it exists a non-zero y_n such that $y_n^\top Q_{\mathrm{X}n-1}^\top \dots Q_{\mathrm{X}1}^\top A_{\mathrm{X}n} = 0$ or transposed

$$A_{\mathrm{X}n}^\top \prod_{i=1}^{n-1} Q_{\mathrm{X}i} \cdot y_n = 0 \ .$$

As all $Q_{\mathrm{X}i}$ have full column rank, by definition,

$$y_k := \prod_{i=k}^{n-1} Q_{\mathrm{X}i} \cdot y_n = Q_{\mathrm{X}k} y_{k+1} \ ,$$

for $k = n-1, n-2, \dots, 1$, are non-trivial and so is y_1. We obtain that

$$A_{\mathrm{X}k}^\top y_1 = A_{\mathrm{X}k}^\top Q_{\mathrm{X}1} Q_{\mathrm{X}2} \dots Q_{\mathrm{X}k-1} y_k = 0 \ , \text{for all } k$$

since $y_k \in \operatorname{im} Q_{\mathrm{X}k} = \ker A_{\mathrm{X}k}^\top Q_{\mathrm{X}1} Q_{\mathrm{X}2} \dots Q_{\mathrm{X}k-1}$. Therefore, $\begin{bmatrix} A_{\mathrm{X}1} & A_{\mathrm{X}2} & \cdots & A_{\mathrm{X}n} \end{bmatrix}$ has no full columns rank and implying that there is a Z-cutset according to Lemma A.6. □

Corollary A.9
Lemma A.8 also holds if type Y groups different branch types $Y_1, Y_2, \dots, Y_m$, for some $m \in \mathbb{N}$, as long as $\mathcal{B}_{\mathrm{Y}} = \bigcup_{j=1}^m \mathcal{B}_{Y_j}$ holds. This is because extending a matrix by columns, that already exist, will not change the row rank. Especially, that means the incidence matrices can double. For example, in case of $m = 2$, replacing A_{Y} with $\begin{bmatrix} A_{\mathrm{Y}1} & A_{\mathrm{Y}1} & A_{\mathrm{Y}2} \end{bmatrix}$ in Lemma A.8 would not change the statement.

A.3 Functional Analysis

Definition A.10
A function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is called Lipschitz continuous with respect to x if and only if there is a constant $L > 0$ such that

$$\|f(x, y) - f(\bar{x}, y)\| \leq L \|x - \bar{x}\| \ , \quad \text{for all } x, \bar{x} \in \mathbb{R}^n \text{ and for all } y \in \mathbb{R}^m \ .$$

The constant L is the function's Lipschitz constant of the first argument.

Definition A.11
With $C^p(\Omega)$ we denote be the space of all p-times continuously differentiable functions that are defined on Ω.

The C^p-norm of a function $f \in C^p(\Omega)$, whose partial derivatives can be continuously extended on the closure $\bar{\Omega}$, is defined by

$$\|f\|_{C^p(\bar{\Omega})} := \max_{|s|\leq p} \|\partial^s f\|_{C^0(\bar{\Omega})} \ ,$$

whereby s is a multiindex and for an arbitrary vector norm $\|\cdot\|$ it is

$$\|f\|_{C^0(\Omega)} := \max_{x\in\Omega} \|f(x)\| \ .$$

Definition A.12 (monotone operators, [Zei90, p. 500])
Let X be real B-spaces, and let $A : X \to X^*$ be an operator. Then:

*(i) A is called **monotone** iff*

$$\langle Au - Av, u - v\rangle \geq 0 \ , \quad \text{for all} \quad u, v \in X \ .$$

*(ii) A is called **strictly monotone** iff*

$$\langle Au - Av, u - v\rangle > 0 \ , \quad \text{for all} \quad u, v \in X \quad \text{with} \quad u \neq v \ .$$

*(iii) A is called **strongly monotone** iff there is an $a > 0$ such that*

$$\langle Au - Av, u - v\rangle \geq a \|u - v\|^2 \ , \quad \text{for all} \quad u, v \in X \ .$$

Remark A.13
From Cauchy-Schwarz inequality, we immediately obtain that a linear operator A which is strongly monotone satisfies, for some $a > 0$,

$$\|Au - Av, u - v\| \, \|u - v\| \geq a \|u - v\|^2$$

and thus, with representation $y = u - v$,

$$\|Ay\| \geq a \|y\| \ , \quad \text{for all } y \in X.$$

Lemma A.14 (Lemma 1 in [Cor+20])
Let $M \in \mathbb{R}^{n\times n}$ be a matrix. Then, the linear function $f(x) := Mx$ is strongly monotone if and only if M is positive definite.

Proof. If $f(x) := Mx$ is strongly monotone we find a constant $a > 0$ such that for all $x \in \mathbb{R}^n$ with $x \neq 0$

$$\langle Mx, x\rangle = \langle f(x) - f(0), x - 0\rangle \geq a\|x - 0\| > 0,$$

that means M is positive definite. Next, we show the opposite direction. Let M be positive definite. We split M into its symmetric and non-symmetric part

$$M = M_s + M_n, \quad M_s = \frac{1}{2}(M + M^\top), \quad M_n = \frac{1}{2}(M - M^\top).$$

Consequently, for all $x \in \mathbb{R}^n$ with $x \neq 0$,

$$\langle M_s x, x\rangle = \langle Mx, x\rangle > 0.$$

Since M_s is symmetric, we find a unitary matrix T and a diagonal matrix D such that $M_s = T^{-1}DT$. We get that

$$0 < \langle M_s x, x\rangle = \langle T^{-1}DTx, x\rangle = \langle DTx, Tx\rangle = \sum_{j=1}^{n} d_{jj} y_j^2 \quad \text{with} \quad y := Tx.$$

Choosing the unit vectors $y := e_i$, we find that $d_{ii} > 0$ for all $i = 1, ..., m$. Defining $a := \min_{i=1,...,m} d_i i$, we see that for all $x \in \mathbb{R}^n$

$$\langle Mx, x\rangle = \langle M_s x, x\rangle \geq \sum_{j=1}^{n} a y_j^2 = a\|Tx\|^2 = a\|x\|^2.$$

Finally, we obtain, for any $x, \bar{x} \in \mathbb{R}^n$

$$\langle f(x) - f(\bar{x}), x - \bar{x}\rangle = \langle M(x - \bar{x}), x - \bar{x}\rangle \geq a\|x - \bar{x}\|^2.$$

□

Lemma A.15 (derivative test)
Let $V \subset \mathbb{R}^n$ open and strictly convex and $f : V \to \mathbb{R}^m$ continuous differentiable. Further, let $\|\cdot\|_V$ and $\|\cdot\|_W$ be vector norms defined on the preimage and image of f, respectively. If the differential of f is bounded by

$$\sup_{z \in V} \|Df(z)\| \leq M,$$

for some $M \geq 0$ using the induced matrix norm, then

$$\|f(x) - f(\bar{x})\|_W \leq M \|x - \bar{x}\|_V \quad \text{for all} \quad x, \bar{x} \in V.$$

Proof. By the vector valued version of the fundamental theorem of calculus we obtain for $x, \bar{x}$

$$f(x) - f(\bar{x}) = \int_0^1 Df(tx + (1-t)x)\,\mathrm{d}t \cdot (x - \bar{x})$$

which is well defined since V is strictly convex and $f \in C^1$. Applying the norm on each side, we obtain

$$\begin{aligned}
\|f(x) - f(\bar{x})\|_W &= \left\| \int_0^1 Df(tx + (1-t)x)\,\mathrm{d}t \cdot (x - \bar{x}) \right\|_W \\
&= \left\| \int_0^1 Df(tx + (1-t)x) \cdot (x - \bar{x})\,\mathrm{d}t \right\|_W \\
&\leq \int_0^1 \|Df(tx + (1-t)x) \cdot (x - \bar{x})\|_W\,\mathrm{d}t \\
&\leq \int_0^1 \|Df(tx + (1-t)x)\|\,\|(x - \bar{x})\|_V\,\mathrm{d}t \\
&\leq \sup_{0\leq t\leq 1} \|Df(tx + (1-t)x)\|\,\|(x - \bar{x})\|_V \qquad \leq \sup_{z\in V} \|Df(z)\|\,\|(x - \bar{x})\|_V .
\end{aligned}$$

The first inequality follows by Cauchy-Schwarz. Since the dot product is a linear continuous operator, the second inequality is an immediate consequence of choosing an induced matrix norm. Finally, if $\sup_{z\in V} \|Df(z)\| \leq M$, the statement follows instantly. □

A.4 Convergence Analysis

This appendix section collects all the proofs, with their exclusively required Lemmata, which have been excluded from the main sections for the sake of readability.

Lemma A.16
Given the Assumptions required by Theorem 6.34 and let every EM devices' branch be part of a V-loop. Then, there always exists a vector norm $\|\cdot\|_*$ so that for the induced matrix norm $\left\|\bar{M}_1\right\|_* + \left\|\bar{M}_2\right\|_* < 1$ with $\bar{M}_1$ and $\bar{M}_2$ defined in (A.8) and (A.9).

Proof. Similar to Lemma 6.28, we use the decoupling in Theorem 3.11 but this time with modifications of Corollary 3.17. We obtain, using the same substitutions as in Theorem 6.34 and while dropping the arguments for the sake of simplicity,

$$\begin{aligned}
\partial_{\mathbf{w}_\mathrm{E}} \tilde{c}_\mathrm{E} &= \begin{bmatrix} 0 \\ 0 \\ -M_\varepsilon^{-1} \end{bmatrix}, \\
\partial_{\mathbf{x}} d_{\mathrm{E-MNA}} &= \begin{bmatrix} M_\varepsilon G_\Gamma \Lambda A_\mathrm{E}^\top & 0 & 0 \end{bmatrix}, \\
T_0 &= \begin{bmatrix} P_\mathrm{C} Q_\mathrm{e} & 0 \\ 0 & \bar{Q}_\mathrm{L} \\ 0 & 0 \end{bmatrix}, \\
\partial_{\mathbf{z}_1} f_0 &= \begin{bmatrix} \partial_{\mathbf{z}_{1l}} f_{0e} & \partial_{\bar{\mathbf{z}}_{1v}} f_{0e} \\ \partial_{\mathbf{z}_{1l}} \bar{f}_{0l} & \partial_{\bar{\mathbf{z}}_{1v}} \bar{f}_{0l} \end{bmatrix} \\
&= \begin{bmatrix} 0 & (Q_\mathrm{e}^\top \hat{C} Q_\mathrm{e})^{-1} Q_\mathrm{e}^\top P_\mathrm{C}^\top [I - \hat{C} P_\mathrm{e} V_\mathrm{C}^\top] A_\mathrm{V} \bar{Q}_\mathrm{V} \\ (\bar{Q}_\mathrm{L}^\top \hat{L}(\bar{\mathbf{y}}_l, \bar{\mathbf{z}}_{3l}) \bar{Q}_\mathrm{L})^{-1} \bar{Q}_\mathrm{L}^\top [I - \hat{L} \bar{P}_\mathrm{L} \bar{V}_\mathrm{L}^\top] \bar{A}_\mathrm{L}^\top & 0 \end{bmatrix},
\end{aligned}$$

$$\partial_{s_{cf}} f_1 = \begin{bmatrix} \partial_{s_{cf}} f_{1l} \\ \partial_{s_{cf}} \bar{f}_{1v} \end{bmatrix} = \begin{bmatrix} 0 \\ -\bar{M}_{1v}(\mathbf{y}_e, \mathbf{z}_{3e}) V_{\mathrm{C}}^\top P_{\mathrm{C}}^\top \end{bmatrix},$$
$$\partial_{s_{cf}} f_0 = \begin{bmatrix} \partial_{s_{cf}} f_{0e} \\ \partial_{s_{cf}} \bar{f}_{0l} \end{bmatrix} = \begin{bmatrix} (Q_{\mathrm{e}}^\top \hat{C} Q_{\mathrm{e}})^{-1} Q_{\mathrm{e}}^\top P_{\mathrm{C}}^\top [I - \hat{C} P_{\mathrm{e}} V_{\mathrm{C}}^\top] \\ 0 \end{bmatrix},$$
$$\partial_{\dot{\mathbf{u}}} s_{cf} = \partial_{\dot{\mathbf{u}}} A_{\mathrm{E}} s_{\mathrm{EM2}} = A_{\mathrm{E}} \Lambda^\top G_\Gamma^\top \begin{bmatrix} M_\varepsilon G & 0 & M_\varepsilon \end{bmatrix}.$$

From this follows

$$\frac{\partial \theta_1}{\partial \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{y}^{[k-1]}} = \partial_{\mathbf{w}_{\mathrm{E}}} \tilde{c}_{\mathrm{E}} \partial_{\mathbf{x}} d_{\mathrm{E-MNA}} T_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -G_\Gamma \Lambda A_{\mathrm{E}}^\top P_{\mathrm{C}} Q_{\mathrm{e}} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ V & 0 \end{bmatrix}$$
$$\begin{aligned} \frac{\partial \theta_2}{\partial \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{u}^{[k]}} &= (\partial_{\mathbf{z}_1} f_0 \partial_{s_{cf}} f_1 + \partial_{s_{cf}} f_0) \partial_{\dot{\mathbf{u}}} s_{cf} \\ &= \Big(\begin{bmatrix} -(Q_{\mathrm{e}}^\top \hat{C} Q_{\mathrm{e}})^{-1} Q_{\mathrm{e}}^\top P_{\mathrm{C}}^\top [I - \hat{C} P_{\mathrm{e}} V_{\mathrm{C}}^\top] A_{\mathrm{V}} \bar{Q}_{\mathrm{V}} \bar{M}_{1v} V_{\mathrm{C}}^\top P_{\mathrm{C}}^\top \\ 0 \end{bmatrix} \\ &\quad + \begin{bmatrix} (Q_{\mathrm{e}}^\top \hat{C} Q_{\mathrm{e}})^{-1} Q_{\mathrm{e}}^\top P_{\mathrm{C}}^\top [I - \hat{C} P_{\mathrm{e}} V_{\mathrm{C}}^\top] \\ 0 \end{bmatrix} \Big) A_{\mathrm{E}} \Lambda^\top G_\Gamma^\top \begin{bmatrix} M_\varepsilon G & 0 & M_\varepsilon \end{bmatrix} \\ &= \begin{bmatrix} UG & 0 & U \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

with

$$V := -G_\Gamma \Lambda A_{\mathrm{E}}^\top P_{\mathrm{C}} Q_{\mathrm{e}}, \tag{A.1a}$$
$$U := (Q_{\mathrm{e}}^\top \hat{C} Q_{\mathrm{e}})^{-1} Q_{\mathrm{e}}^\top P_{\mathrm{C}}^\top [I - \hat{C} P_{\mathrm{e}} V_{\mathrm{C}}^\top][I - A_{\mathrm{V}} \bar{Q}_{\mathrm{V}} \bar{M}_{1v} V_{\mathrm{C}}^\top P_{\mathrm{C}}^\top] A_{\mathrm{E}} \Lambda^\top G_\Gamma^\top M_\varepsilon. \tag{A.1b}$$

Taking the EM devices as substitutions for the mock elements, from Lemma 6.33 we obtain that $Q_{\mathrm{e}}^\top P_{\mathrm{C}}^\top A_{\mathrm{E}} = 0$ and, consequently,

$$\bar{M}_1 := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \bar{M}_2 := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ UG & 0 & U & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

In order to find a vector norm $\|\cdot\|_*$ so that for the induced matrix norm holds $\|\bar{M}_1\|_* + \|\bar{M}_2\|_* < 1$ we have to exploit the block matrix structures. Let $v = (v_1, v_2, v_3, v_4, v_5)$ be in the pre-image of $\bar{M}_2$ and $\bar{M}_2$. We choose a vector norm, e. g. the max-norm, and define

$$\varepsilon_{UG} := \sup_{v_1 \neq 0} \frac{\|UGv_1\|_\infty}{\|v_1\|_\infty}, \qquad \varepsilon_U := \sup_{v_3 \neq 0} \frac{\|Uv_3\|_\infty}{\|v_3\|_\infty}.$$

With these constants, we define the three custom norms

$$\|v_1\|_{1*} := 4\varepsilon_{UG} \|v_1\|_\infty, \qquad \|v_3\|_{3*} := 4\varepsilon_U \|v_3\|_\infty$$

and finally the one acting on whole v by

$$\|v\|_* = \left\| \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} \right\|_* := \|v_1\|_{1*} + \|v_2\|_\infty + \|v_3\|_{3*} + \|v_4\|_\infty + \|v_5\|_\infty . \tag{A.2}$$

For the induced matrix norm of $\|\cdot\|_*$ holds

$$\left\|\bar{M}_1\right\|_* = \sup_{v \neq 0} \frac{\left\|\bar{M}_1 v\right\|_*}{\|v\|_*} = 0$$

and

$$\begin{aligned}
\left\|\bar{M}_2\right\|_* &= \sup_{v \neq 0} \frac{\left\|\bar{M}_2 v\right\|_*}{\|v\|_*} \\
&= \sup_{v \neq 0} \frac{\|UGv_1 + Uv_3\|_\infty}{\|v_1\|_{1*} + \|v_2\|_\infty + \|v_3\|_{3*} + \|v_4\|_\infty + \|v_5\|_\infty} \\
&\leq \sup_{v \neq 0} \frac{\|UGv_1\|_\infty + \|Uv_3\|_4}{\|v_1\|_{1*} + \|v_3\|_{3*}} \\
&\leq \sup_{v \neq 0} \frac{\|UGv_1\|_\infty}{\|v_1\|_{1*}} + \frac{\|Uv_3\|_\infty}{\|v_3\|_{3*}} \\
&= \sup_{v \neq 0} \frac{1}{4\varepsilon_{UG}} \frac{\|UGv_1\|_\infty}{\|v_1\|_\infty} + \frac{1}{4\varepsilon_U} \frac{\|Uv_3\|_\infty}{\|v_3\|_\infty} = \frac{1}{4} + \frac{1}{4}
\end{aligned}$$

In conclusion, it is $\left\|\bar{M}_1\right\|_* + \left\|\bar{M}_2\right\|_* \leq \frac{1}{2} < 1$. □

Proof of Theorem 6.34. The proof proceeds similar to the one of Theorem 6.29. In addition, Assumption 3.16 is naturally fulfilled as of Remark 5.9 and, therefore, especially 3.10. Consequently, all the required assumptions for Theorem 3.11 are met. Hence, for any iteration counter k, we can decouple (6.30b) by making use of Corollary 3.17 while setting $f := f_{\text{MNA2}}$, $d := d_{\text{MNA2}}$, $A_{\text{C}} := A_{\mathcal{C}}, A_{\text{R}} := A_{\mathcal{R}}$ and

$$s_c(\mathbf{i}_{\text{M}}) = A_{\text{M}}\mathbf{i}_{\text{M}} = A_{\text{E}} s_{\text{E}}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u}) =: s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}, \mathbf{u})$$

yielding the equivalent system

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]} = f_0(\mathbf{y}^{[k]}, \mathbf{z}_1^{[k]}, \mathbf{z}_2^{[k]}, \mathbf{z}_3^{[k]}, s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]})), \tag{A.3a}$$

$$\mathbf{z}_1^{[k]} = M_1(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}) \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k]} + f_1(\mathbf{y}^{[k]}, \mathbf{z}_2^{[k]}, \mathbf{z}_3^{[k]}, s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]})), \tag{A.3b}$$

$$\mathbf{z}_2^{[k]} = f_2(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}, s_v, s_i), \tag{A.3c}$$

$$\mathbf{z}_3^{[k]} = M_3 \begin{pmatrix} s_i \\ s_v \end{pmatrix} \tag{A.3d}$$

with the splitting $\mathbf{x}^{[k]} = T_0\mathbf{y}^{[k]} + T_1\mathbf{z}_1^{[k]} + T_2\mathbf{z}_2^{[k]} + T_3\mathbf{z}_3^{[k]}$ for transformation matrices

$$T_0 = \begin{bmatrix} P_\mathrm{C}Q_\mathrm{e} & 0 \\ 0 & \bar{Q}_\mathrm{L} \\ 0 & 0 \end{bmatrix},\ T_1 = \begin{bmatrix} Q_\mathrm{C}Q_\mathrm{V}Q_\mathrm{R} & 0 \\ 0 & 0 \\ 0 & \bar{Q}_\mathrm{V} \end{bmatrix},\ T_2 = \begin{bmatrix} Q_\mathrm{C}P_\mathrm{V} & Q_\mathrm{C}Q_\mathrm{V}P_\mathrm{R} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{P}_\mathrm{V} \end{bmatrix},\ T_3 = \begin{bmatrix} P_\mathrm{C}P_\mathrm{e} & 0 \\ 0 & \bar{P}_\mathrm{L} \\ 0 & 0 \end{bmatrix}.$$

Resolving the variables in (A.3) by inserting them into each other yields

$$\begin{aligned}
\mathbf{z}_3^{[k]} &= M_3\begin{pmatrix} s_i \\ s_v \end{pmatrix} =: m_3(s_v, s_i) \\
\mathbf{z}_2^{[k]} &= f_2(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}, s_v, s_i) \\
&= f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i) \\
\mathbf{z}_1^{[k]} &= M_1(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]})\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k]} + f_1(\mathbf{y}^{[k]}, \mathbf{z}_2^{[k]}, \mathbf{z}_3^{[k]}, s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]})) \\
&= M_1(\mathbf{y}^{[k]}, m_3(s_v, s_i))m_3(\dot{s}_v, \dot{s}_i) + f_1(\mathbf{y}^{[k]}, f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i), \\
&\quad m_3(s_v, s_i), s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}))
\end{aligned}$$

and for (A.3a) we obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]} &= f_0(\mathbf{y}^{[k]}, \big[M_1(\mathbf{y}^{[k]}, m_3(s_v, s_i))m_3(\dot{s}_v, \dot{s}_i) \\
&\qquad + f_1(\mathbf{y}^{[k]}, f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i), m_3(s_v, s_i), s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}))\big], \\
&\qquad f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i), m_3(s_v, s_i), s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]})) \\
&=: \theta_2(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}, \mathbf{y}^{[k]}, t)
\end{aligned} \tag{A.4}$$

Further, we notice that $A_\mathrm{E}^\top Q_C = 0$ since capacitances and resistances from the EM coupling equation (5.17) were shifted in into the MNA, that is $\frac{\mathrm{d}}{\mathrm{d}t}q_\mathrm{E}(A_\mathrm{E}^\top\mathbf{e}) + g_\mathrm{E}(A_\mathrm{E}^\top\mathbf{e})$ with strongly monotone q_E according to Lemma 5.7. In other words, for the basis matrix Q_C holds $\operatorname{im} Q_\mathrm{C} = \ker A_\mathcal{C}^\top$ and, especially, $\operatorname{im} Q_\mathrm{C}^\top \subset \ker A_\mathrm{E}$. Therefore and by definitions of linear $d_{\mathrm{E-MNA}}$ and c_E in (5.9), it is

$$\begin{aligned}
d_{\mathrm{E-MNA}}(T_1z_1) &= \begin{bmatrix} M_\varepsilon G_\Gamma \Lambda A_\mathrm{E}^\top & 0 & 0 \end{bmatrix}\begin{bmatrix} Q_\mathrm{C}Q_\mathrm{V}Q_\mathrm{R} & 0 \\ 0 & 0 \\ 0 & \bar{Q}_\mathrm{V} \end{bmatrix} z_1 = 0, \\
d_{\mathrm{E-MNA}}(T_2z_2) &= \begin{bmatrix} M_\varepsilon G_\Gamma \Lambda A_\mathrm{E}^\top & 0 & 0 \end{bmatrix}\begin{bmatrix} Q_\mathrm{C}P_\mathrm{V} & Q_\mathrm{C}Q_\mathrm{V}P_\mathrm{R} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{P}_\mathrm{V} \end{bmatrix} z_2 = 0, \\
c_\mathrm{E}(w, T_1z_1) &= -\begin{pmatrix} 0 \\ w \\ 0 \end{pmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ M_\sigma G_\Gamma \Lambda A_\mathrm{E}^\top & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} Q_\mathrm{C}Q_\mathrm{V}Q_\mathrm{R} & 0 \\ 0 & 0 \\ 0 & \bar{Q}_\mathrm{V} \end{bmatrix} z_1 = c_\mathrm{E}(w, 0), \\
c_\mathrm{E}(w, T_2z_2) &= -\begin{pmatrix} 0 \\ w \\ 0 \end{pmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ M_\sigma G_\Gamma \Lambda A_\mathrm{E}^\top & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} Q_\mathrm{C}P_\mathrm{V} & Q_\mathrm{C}Q_\mathrm{V}P_\mathrm{R} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{P}_\mathrm{V} \end{bmatrix} z_2 = c_\mathrm{E}(w, 0).
\end{aligned}$$

Exploiting these observations, inserting the variable splitting with the proper iteration counter into (6.54), which is equivalent to (6.30a), yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} =& -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(T_0\mathbf{y}^{[k-1]} + T_1\mathbf{z}_1^{[k-1]} + T_2\mathbf{z}_2^{[k-1]} + T_3\mathbf{z}_3^{[k-1]}),\\ & T_0\mathbf{y}^{[k-1]} + T_1\mathbf{z}_1^{[k-1]} + T_2\mathbf{z}_2^{[k-1]} + T_3\mathbf{z}_3^{[k-1]})\\ =& -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(T_0\mathbf{y}^{[k-1]} + T_3\mathbf{z}_3^{[k-1]}), T_0\mathbf{y}^{[k-1]} + T_3\mathbf{z}_3^{[k-1]})\\ =& -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(d_{\mathrm{E-MNA}}(T_0\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]} + T_3\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k-1]}),\\ & T_0\mathbf{y}^{[k-1]} + T_3\mathbf{z}_3^{[k-1]}).\end{aligned} \tag{A.5}$$

Again, we insert the remaining algebraic expressions from the partial decoupling, i. e., with proper iteration counter, $\mathbf{z}_3^{[k-1]} = m_3(s_i, s_v)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k-1]} = m_3(\dot{s}_i, \dot{s}_v)$. Thus, we can rewrite (A.5) as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} =& -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(d_{\mathrm{E-MNA}}(T_0\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]} + T_3 m_3(\dot{s}_i, \dot{s}_v),\\ & T_0\mathbf{y}^{[k-1]} + T_3 m_3(s_i, s_v))\\ =:& \ \theta_1(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{y}^{[k-1]}, t).\end{aligned} \tag{A.6}$$

From (A.6) and (A.4) we obtain a waveform relaxation scheme of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}(t) = \theta(\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}(t), \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}(t), x^{[k]}(t), x^{[k-1]}(t), t), \quad x^{[k]}(t_0) = (u_0, y_0) \tag{A.7}$$

with $\theta = (\theta_1, \theta_2)$, for $x^{[k]} = (\mathbf{u}^{[k]}, \mathbf{y}^{[k]})$, iteration counter $k \in \mathbb{N}$ and initial guess $x^{[0]} = (\cdot, \mathbf{y}^{[0]})$ so that $x^{[0]}(t_0) = (\cdot, y_0)$. Note that the previous dependency of t was dropped due to readability. Now, θ is continuous and Lipschitz continuous in the first four arguments by construction in Theorem 3.11, by Assumption 3.12 and by Lemma 5.5. Respectively, we denote the Lipschitz-constants, which are independent of t, by c_1, c_2, d_1, d_2. In order to apply Theorem 6.21, by setting $G = \theta$, we first have to show that there exists a vector norm so that $c_1 + c_2 < 1$. Again, using the derivative test, see Lemma A.15, consider the still bounded

$$\bar{M}_1 := \frac{\partial\theta}{\partial\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}} = \begin{bmatrix}\frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}} & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]}}\\ \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}} & \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]}}\end{bmatrix} = \begin{bmatrix}0 & 0\\ \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}} & 0\end{bmatrix}, \tag{A.8}$$

$$\bar{M}_2 := \frac{\partial\theta}{\partial\frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}} = \begin{bmatrix}\frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}}\\ \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} & \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}}\end{bmatrix} = \begin{bmatrix}0 & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}}\\ 0 & 0\end{bmatrix} \tag{A.9}$$

with

$$\begin{aligned}\frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}} &= [\partial_{\mathbf{z}_1} f_0 \partial_{s_{cf}} f_1 + \partial_{s_{cf}} f_0]\partial_{\dot{\mathbf{u}}} s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}),\\ \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}} &= \partial_{\mathbf{w}_{\mathrm{E}}}\tilde{c}_{\mathrm{E}}\partial_{\mathbf{x}} d_{\mathrm{E-MNA}} T_0.\end{aligned}$$

Note that especially due to the Lipschitz-continuity of θ in the first two arguments, these partial derivatives are bounded. As of Lemma A.16, there exists a vector norm $\|\cdot\|$ so that the induced matrix norm yields $\|\bar{M}_1\| + \|\bar{M}_2\| < 1$. Hence, we can apply Lemma A.15 twice guaranteeing that $c_1 + c_2 < 1$. From Theorem 6.21 we deduce that the sequence $x^{[k]}$ in (6.60) converges to $x^* = (\mathbf{u}^*, \mathbf{y}^*) \in C^1(\mathcal{I}, \mathbb{R}^n)$ which is the unique solution of

$$\frac{\mathrm{d}}{\mathrm{d}t}x = \theta(\frac{\mathrm{d}}{\mathrm{d}t}x, \frac{\mathrm{d}}{\mathrm{d}t}x, x, x, t), \quad x(t_0) = (u_0, y_0).$$

Then, we exploit the (Lipschitz) continuities of f_1, f_2 as well as the boundedness of M_1 and obtain, for the limits of (A.3b)-(A.3d), as we assumed sufficient smooth source functions,

$$\begin{aligned}
\lim_{k\to\infty} \mathbf{z}_1^{[k]} &= \lim_{k\to\infty} M_1(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]})\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k]} + f_1(\mathbf{y}^{[k]}, \mathbf{z}_2^{[k]}, \mathbf{z}_3^{[k]}, s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]})) \\
&= M_1(\mathbf{y}^*, \mathbf{z}_3^*)\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^* + f_1(\mathbf{y}^*, \mathbf{z}_2^*, \mathbf{z}_3^*, s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^*, \mathbf{u}^*)) = \mathbf{z}_1^*, \\
\lim_{k\to\infty} \mathbf{z}_2^{[k]} &= \lim_{k\to\infty} f_2(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}, s_v, s_i) \\
&= f_2(\mathbf{y}^*, \mathbf{z}_3^*, s_v, s_i) = \mathbf{z}_2^*
\end{aligned}$$

where $\mathbf{z}_3^{[k]} = m_3(s_v, s_i) = \mathbf{z}_3^*$. Consequently,

$$\lim_{k\to\infty} \mathbf{x}^{[k]} = \lim_{k\to\infty} T_0\mathbf{y}^{[k]} + T_1\mathbf{z}_1^{[k]} + T_2\mathbf{z}_2^{[k]} + T_3\mathbf{z}_3^{[k]} = T_0\mathbf{y}^* + T_1\mathbf{z}_1^* + T_2\mathbf{z}_2^* + T_3\mathbf{z}_3^* = \mathbf{x}^*$$

converges in $C^0(\mathcal{I}, \mathbb{R}^n)$ with the according C^0-norm. □

Proof of Theorem 6.36. The proof's start is analogue to the one of Theorem 6.29 but for the following changes:

$$\begin{aligned}
\mathbf{z}_3^{[k]} &= M_3\begin{pmatrix} s_i \\ s_v \end{pmatrix} =: m_3(s_v, s_i) \\
\mathbf{z}_2^{[k]} &= f_2(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}, s_v, s_i, s_{cf}(\mathbf{u}^{[k-1]})) \\
&= f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i, s_{cf}(\mathbf{u}^{[k-1]})) \\
&=: f_2^{[k]} \\
\mathbf{z}_1^{[k]} &= M_1(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]})\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k]} + f_1(\mathbf{y}^{[k]}, \mathbf{z}_2^{[k]}, \mathbf{z}_3^{[k]}, s_i, s_{cf}(\mathbf{u}^{[k-1]})) \\
&= M_1(\mathbf{y}^{[k]}, m_3(s_v, s_i))m_3(\dot{s}_v, \dot{s}_i) + f_1(\mathbf{y}^{[k]}, f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i, s_{cf}(\mathbf{u}^{[k-1]})), \\
&\quad m_3(s_v, s_i), s_i, s_{cf}(\mathbf{u}^{[k-1]}))
\end{aligned}$$

with derivatives

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k]} &= \frac{\mathrm{d}}{\mathrm{d}t}m_3(s_v, s_i), \\
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_2^{[k]} &= \frac{\partial}{\partial \mathbf{y}}f_2^{[k]}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]} + \frac{\partial}{\partial \mathbf{z}_3}f_2^{[k]}\frac{\mathrm{d}}{\mathrm{d}t}m_3(s_v, s_i) \\
&\quad + \frac{\partial}{\partial s_v}f_2^{[k]}\frac{\mathrm{d}}{\mathrm{d}t}s_v + \frac{\partial}{\partial s_i}f_2^{[k]}\frac{\mathrm{d}}{\mathrm{d}t}s_i + \frac{\partial}{\partial s_{cf}}f_2^{[k]}\frac{\partial}{\partial \mathbf{u}}s_{cf}(\mathbf{u}^{[k-1]})\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}.
\end{aligned}$$

Yielding

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]} &= f_0(\mathbf{y}^{[k]}, \big[M_1(\mathbf{y}^{[k]}, m_3(s_v,s_i))\frac{\mathrm{d}}{\mathrm{d}t}m_3(s_v,s_i)\\ &\quad + f_1(\mathbf{y}^{[k]}, f_2(\mathbf{y}^{[k]}, m_3(s_v,s_i), s_v, s_i, s_{cf}(\mathbf{u}^{[k-1]})), m_3(s_v,s_i), s_i, s_{cf}(\mathbf{u}^{[k-1]}))\big],\\ &\quad f_2(\mathbf{y}^{[k]}, m_3(s_v,s_i), s_v, s_i, s_{cf}(\mathbf{u}^{[k-1]})), m_3(s_v,s_i), s_i, s_{cf}(\mathbf{u}^{[k-1]}))\\ &=: \theta_2(\mathbf{y}^{[k]}, \mathbf{u}^{[k-1]}, t)\end{aligned} \tag{A.10}$$

and

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} &= -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t}d_{\mathrm{E-MNA}}(T_0\mathbf{y}^{[k-1]} + T_1\mathbf{z}_1^{[k-1]} + T_2\mathbf{z}_2^{[k-1]} + T_3\mathbf{z}_3^{[k-1]}),\\ &\quad T_0\mathbf{y}^{[k-1]} + T_1\mathbf{z}_1^{[k-1]} + T_2\mathbf{z}_2^{[k-1]} + T_3\mathbf{z}_3^{[k-1]})\\ &= -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(d_{\mathrm{E-MNA}}(T_0\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]} + T_2\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_2^{[k-1]} + T_3\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k-1]}),\\ &\quad T_0\mathbf{y}^{[k-1]} + T_2\mathbf{z}_2^{[k-1]} + T_3\mathbf{z}_3^{[k-1]})\\ &= -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(d_{\mathrm{E-MNA}}(T_0\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]} + T_2\big[\frac{\partial}{\partial\mathbf{y}}f_2^{[k-1]}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}\\ &\quad + \frac{\partial}{\partial\mathbf{z}_3}f_2^{[k-1]}\frac{\mathrm{d}}{\mathrm{d}t}m_3(s_v,s_i) + \frac{\partial}{\partial s_v}f_2^{[k-1]}\frac{\mathrm{d}}{\mathrm{d}t}s_v + \frac{\partial}{\partial s_i}f_2^{[k-1]}\frac{\mathrm{d}}{\mathrm{d}t}s_i\\ &\quad + \frac{\partial}{\partial s_{cf}}f_2^{[k-1]}\frac{\partial}{\partial\mathbf{u}}s_{cf}(\mathbf{u}^{[k-2]})\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-2]}\big] + T_3\frac{\mathrm{d}}{\mathrm{d}t}m_3(s_v,s_i)),\\ &\quad T_0\mathbf{y}^{[k-1]} + T_2f_2^{[k-1]} + T_3m_3(s_v,s_i))\\ &=: \theta_1(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}, \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-2]}, \mathbf{u}^{[k]}, \mathbf{y}^{[k-1]}, \mathbf{u}^{[k-2]}, t).\end{aligned} \tag{A.11}$$

From (A.11) and (A.10) we have obtained a waveform relaxation scheme of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]} = \theta(\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}, \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}, \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-2]}, x^{[k]}, x^{[k-1]}, x^{[k-2]}, t), \quad x^{[k]}(t_0) = (u_0, y_0) \tag{A.12}$$

with $\theta = (\theta_1, \theta_2)$, for $x^{[k]} = (\mathbf{u}^{[k]}, \mathbf{y}^{[k]})$, iteration counter $k \in \mathbb{N}$ and initial guess $x^{[0]} = (\mathbf{u}^{[0]}, \mathbf{y}^{[0]})$ so that $x^{[0]}(t_0) = (u_0, y_0)$. Note that the previous dependency of t of the unknowns was dropped due to readability.

Again, θ is continuous and Lipschitz continuous in the first six arguments by construction in Theorem 3.11, by Assumption 3.12 and by Lemma 5.5 with time independent Lipschitz-constants $c_1, c_2, c_3, d_1, d_2, d_3$. Consider the thus bounded partial derivatives

$$\bar{M}_1 := \frac{\partial\theta}{\partial\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}} = \begin{bmatrix}\frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}} & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]}}\\ \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}} & \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]}}\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix},$$

$$\bar{M}_2 := \frac{\partial\theta}{\partial\frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}} = \begin{bmatrix}\frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}}\\ \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} & \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}}\end{bmatrix} = \begin{bmatrix}0 & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}}\\ 0 & 0\end{bmatrix} = \begin{bmatrix}0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & V & W\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{bmatrix},$$

$$\bar{M}_3 := \frac{\partial\theta}{\partial \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-2]}} = \begin{bmatrix} \frac{\partial\theta_1}{\partial \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-2]}} & \frac{\partial\theta_1}{\partial \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-2]}} \\ \frac{\partial\theta_2}{\partial \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-2]}} & \frac{\partial\theta_2}{\partial \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-2]}} \end{bmatrix} = \begin{bmatrix} \frac{\partial\theta_1}{\partial \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-2]}} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & U & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

with U, V and W as in (6.55). Hence, with the norm $\|\cdot\|_*$ in (6.56) we already found a norm such that $\left\|\bar{M}_1\right\|_* + \left\|\bar{M}_2\right\|_* + \left\|\bar{M}_3\right\|_* < 1$. Finally we apply Lemma A.15 multiple times, then Theorem 6.19 and conclude similar as in proof of Theorem 6.29. □

Proof of Theorem 6.39. The proof's start is analogue to the one of Theorem 6.34 but for the following:

$$\begin{aligned}
\mathbf{z}_3^{[k]} &= M_3 \begin{pmatrix} s_i \\ s_v \end{pmatrix} =: m_3(s_v, s_i) \\
\mathbf{z}_2^{[k]} &= f_2(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}, s_v, s_i) \\
&= f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i) \\
\mathbf{z}_1^{[k]} &= M_1(\mathbf{y}^{[k]}, \mathbf{z}_3^{[k]}) \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k]} + f_1(\mathbf{y}^{[k]}, \mathbf{z}_2^{[k]}, \mathbf{z}_3^{[k]}, s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}, \mathbf{u}^{[k-1]})) \\
&= M_1(\mathbf{y}^{[k]}, m_3(s_v, s_i)) m_3(\dot{s}_v, \dot{s}_i) + f_1(\mathbf{y}^{[k]}, f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i), \\
&\quad m_3(s_v, s_i), s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}, \mathbf{u}^{[k-1]}))
\end{aligned}$$

yielding

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]} &= f_0(\mathbf{y}^{[k]}, \big[M_1(\mathbf{y}^{[k]}, m_3(s_v, s_i)) m_3(\dot{s}_v, \dot{s}_i) \\
&\quad + f_1(\mathbf{y}^{[k]}, f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i), m_3(s_v, s_i), s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}, \mathbf{u}^{[k-1]}))\big], \\
&\quad f_2(\mathbf{y}^{[k]}, m_3(s_v, s_i), s_v, s_i), m_3(s_v, s_i), s_i, s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}, \mathbf{u}^{[k-1]})) \\
&=: \theta_2(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}, \mathbf{y}^{[k]}, \mathbf{u}^{[k-1]}, t).
\end{aligned} \tag{A.13}$$

With $\mathbf{z}_3^{[k-1]} = m_3(s_i, s_v)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k-1]} = m_3(\dot{s}_i, \dot{s}_v)$ we further obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]} &= -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(T_0\mathbf{y}^{[k-1]} + T_1\mathbf{z}_1^{[k-1]} + T_2\mathbf{z}_2^{[k-1]} + T_3\mathbf{z}_3^{[k-1]}), \\
&\quad T_0\mathbf{y}^{[k-1]} + T_1\mathbf{z}_1^{[k-1]} + T_2\mathbf{z}_2^{[k-1]} + T_3\mathbf{z}_3^{[k-1]}) \\
&= -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathrm{E-MNA}}(T_0\mathbf{y}^{[k-1]} + T_3\mathbf{z}_3^{[k-1]}), T_0\mathbf{y}^{[k-1]} + T_3\mathbf{z}_3^{[k-1]}) \\
&= -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(d_{\mathrm{E-MNA}}(T_0\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]} + T_3\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z}_3^{[k-1]}), \\
&\quad T_0\mathbf{y}^{[k-1]} + T_3\mathbf{z}_3^{[k-1]}) \\
&= -\tilde{b}_{\mathrm{E}}(\mathbf{u}^{[k]}) + \tilde{c}_{\mathrm{E}}(d_{\mathrm{E-MNA}}(T_0\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]} + T_3 m_3(\dot{s}_i, \dot{s}_v),
\end{aligned}$$

$$
\begin{aligned}
&T_0\mathbf{y}^{[k-1]} + T_3 m_3(s_i, s_v))\\
&=: \theta_1(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}, \mathbf{u}^{[k]}, \mathbf{y}^{[k-1]}, t).
\end{aligned} \tag{A.14}
$$

From (A.14) and (A.13) we obtain a waveform relaxation scheme of the form

$$
\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]} = \theta(\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}, \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}, \frac{\mathrm{d}}{\mathrm{d}t}x^{[k-2]}, x^{[k]}, x^{[k-1]}, x^{[k-2]}, t), \quad x^{[k]}(t_0) = (u_0, y_0) \tag{A.15}
$$

with $\theta = (\theta_1, \theta_2)$, for $x^{[k]} = (\mathbf{u}^{[k]}, \mathbf{y}^{[k]})$, iteration counter $k \in \mathbb{N}$ and initial guess $x^{[0]} = (\mathbf{u}^{[0]}, \mathbf{y}^{[0]})$ so that $x^{[0]}(t_0) = (u_0, y_0)$. Note that the previous dependency of t of the unknowns was dropped due to readability.
Again, θ is continuous and Lipschitz continuous in the first four arguments by construction in Theorem 3.11, by Assumption 3.12 and by Lemma 5.5 and has time independent Lipschitz-constants c_1, c_2, d_1, d_2. Consider the hence bounded partial derivatives

$$
\begin{aligned}
\bar{M}_1 &:= \frac{\partial\theta}{\partial\frac{\mathrm{d}}{\mathrm{d}t}x^{[k]}} = \begin{bmatrix} \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}} & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]}} \\ \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}} & \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k]}} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},\\
\bar{M}_2 &:= \frac{\partial\theta}{\partial\frac{\mathrm{d}}{\mathrm{d}t}x^{[k-1]}} = \begin{bmatrix} \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}} \\ \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} & \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}} \end{bmatrix} = \begin{bmatrix} 0 & \frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}} \\ \frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k-1]}} & 0 \end{bmatrix}
\end{aligned} \tag{A.16}
$$

with

$$
\begin{aligned}
\frac{\partial\theta_2}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}} &= [\partial_{\mathbf{z}_1} f_0 \partial_{s_{cf}} f_1 + \partial_{s_{cf}} f_0]\partial_{\dot{\mathbf{u}}} s_{cf}(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^{[k]}, \mathbf{u}^{[k]}),\\
\frac{\partial\theta_1}{\partial\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}^{[k-1]}} &= \partial_{\mathbf{w}_\mathrm{E}} \tilde{c}_\mathrm{E} \partial_{\mathbf{x}} d_{\mathrm{E-MNA}} T_0.
\end{aligned}
$$

Since we assumed that for a certain vector norm it holds $\left\|\bar{M}_2\right\|_* < 1$, we apply Lemma A.15, then Theorem 6.21 and conclude similar as in proof of Theorem 6.34. □

Bibliography

[AE01] H. Amann and J. Escher. "Analysis II. Grundstudium Mathematik". In: *Analysis. Birkhäuser Basel* 1 (2001).

[AG01] M. Arnold and M. Günther. "Preconditioned Dynamic Iteration for Coupled Differential-Algebraic Systems". In: *BIT Numerical Mathematics* 41.1 (Jan. 2001), pp. 1–25.

[AGR14] M. Al-Khaleel, M. J. Gander, and A. E. Ruehli. "A mathematical analysis of optimized waveform relaxation for a small RC circuit". In: *Applied Numerical Mathematics* 75 (2014), pp. 61–76.

[Alì+12] G. Alì, A. Bartel, M. Brunk, and S. Schöps. "A Convergent Iteration Scheme for Semiconductor/Circuit Coupled Problems." In: *Scientific Computing in Electrical Engineering SCEE 2010. Mathematics in Industry.* Ed. by B. Michielsen and J. Poirier. Vol. 16. Springer, Berlin, Heidelberg, 2012, pp. 104–111.

[Alì+13] G. Alì, N. Banagaaya, W. H. A. Schilders, and C. Tischendorf. "Index-aware model order reduction for linear index-2 DAEs with constant coefficients". In: *SIAM Journal on Scientific Computing* 35.3 (2013), A1487–A1510.

[AR03] A. Alonso Rodríguez and M. Raffetto. "Unique solvability for electromagnetic boundary value problems in the presence of partly lossy inhomogeneous anisotropic media and mixed boundary conditions". In: *Mathematical Models and Methods in Applied Sciences* 13.04 (2003), pp. 597–611.

[Arg92] P. N. Argyres. "Quantum kinetic equations for electrons in high electric and phonon fields". In: *Physics Letters A* 171.5-6 (1992), pp. 373–379.

[Arn+04] A. Arnold, J. L. López, P. A. Markowich, and J. Soler. "An analysis of quantum Fokker-Planck models: A Wigner function approach". In: *Revista Matemática Iberoamericana* 20.3 (2004), pp. 771–814.

[AV10] A. Alonso Rodríguez and A. Valli. *Eddy Current Approximation of Maxwell Equations.* Springer Milan, 2010.

[AV98] A. Alonso and A. Valli. "Unique solvability for high-frequency heterogeneous time-harmonic Maxwell equations via the Fredholm alternative theory". In: *Mathematical Methods in the Applied Sciences* 21.6 (1998), pp. 463–477.

[BAN00] A. Buffa, H. Ammari, and J.-C. Nédélec. "A justification of eddy currents model for the Maxwell equations". In: *SIAM Journal on Applied Mathematics* 60.5 (2000), pp. 1805–1823.

[Ban10] N. Banagaaya. "The EV-formulation and an investigation of eigenvalues for electromagnetic problems". MA thesis. Eindhoven University of Technology, Aug. 2010.

[Bar+13] A. Bartel, M. Brunk, M. Günther, and S. Schöps. "Dynamic Iteration for Coupled Problems of Electric Circuits and Distributed Devices". In: *SIAM Journal on Scientific Computing* 35.2 (2013), B315–B335.

[Bar+16] A. Bartel, M. Clemens, M. Günther, and E. J. W. ter Maten. *Scientific computing in electrical engineering*. Springer, 2016.

[Bau+13] S. Baumanns, L. Jansen, M. Selva-Soto, and C. Tischendorf. "PDAE Analysis for Coupled Circuit Device Simulation with Finite and Mixed-Finite Elements". Institut für Mathematik, Humboldt-Universität zu Berlin (ISSN 0863-0976), 2013.

[Bau12] S. Baumanns. *Coupled Electromagnetic Field / Circuit Simulation. Modeling and Numerical Analysis*. Logos Verlag, Berlin, 2012.

[Bau72] J. Baumgarte. "Stabilization of constraints and integrals of motion in dynamical systems". In: *Computer methods in applied mechanics and engineering* 1.1 (1972), pp. 1–16.

[BBS14] A. Bartel, M. Brunk, and S. Schöps. "On the convergence rate of dynamic iteration for coupled problems with multiple subsystems". In: *Journal of Computational and Applied Mathematics* 262 (2014), pp. 14–24.

[BCP89] K. E. Brenan, S. L. Campbell, and L. R. Petzold. *The Numerical Solution of Initial Value Problems in Ordinary Differential–Algebraic Equations*. North Holland Publishing Co., New York, 1989.

[Bed93] G. Bedrosian. "A new method for coupling finite element field solutions with external circuits and kinematics". In: *IEEE Transactions on Magnetics* 29.2 (1993), pp. 1664–1668.

[Ben06] G. Benderskaya. "Numerical Methods for Transient Field-Circuit Coupled Simulations Based on the Finite Integration Technique and a Mixed Circuit Formulation". PhD thesis. Technische Universität Darmstadt, 2006.

[BF15] P. Benner and L. Feng. "Model order reduction for coupled problems". In: *Appl Comput Math Int J* 14.1 (2015), pp. 3–22.

[BGL05] M. Benzi, G. H. Golub, and J. Liesen. "Numerical solution of saddle point problems". In: *Acta numerica* 14 (2005), p. 1.

[Bír+90] O. Bíró, K. Preis, W. Renhart, K. R. Richter, and G. Vrisk. "Performance of different vector potential formulations in solving multiply connected 3-D eddy current problems". In: *IEEE Transactions on Magnetics* 26.2 (1990), pp. 438–441.

[Bit+18] K. Bittner, H. G. Brachtendorf, W. Schoenmaker, C. Strohm, and C. Tischendorf. "Coupled Circuit Device Simulation". In: *Scientific Computing in Electrical Engineering: SCEE 2016, St. Wolfgang, Austria, October 2016*. Springer. 2018, pp. 69–77.

[BK00] A. Bossavit and L. Kettunen. "Yee-like schemes on staggered cellular grids: A synthesis between FIT and FEM approaches". In: *IEEE Transactions on Magnetics* 36.4 (2000), pp. 861–867.

[Blo+11] T. Blochwitz, M. Otter, M. Arnold, C. Bausch, C. Clauss, H. Elmqvist, A. Junghanns, J. Mauss, M. Monteiro, T. Neidhold, D. Neumerkel, H. Olosson, J.-V. Peetz, and S. Wolf. "The functional mockup interface for tool independent exchange of simulation models". In: *Proceedings of the 8th International Modelica Conference*. Linköping University Press. 2011, pp. 105–114.

[Bos88] A. Bossavit. "Whitney forms: A class of finite elements for three-dimensional computations in electromagnetism". In: *IEE Proceedings A (Physical Science, Measurement and Instrumentation, Management and Education, Reviews)* 135.8 (1988), pp. 493–500.

[Bos91] A. Bossavit. "Differential geometry for the student of numerical methods in Electromagnetism". In: (1991).

[Bos98] A. Bossavit. *Computational Electromagnetism: Variational Formulations, Complementarity, Edge Elements*. Academic Press, 1998.

[Bos99] A. Bossavit. "Comqutational electromagnetism and geometry:(2): Network constitutive laws". In: *Journal of the Japan Society of Applied Electromagnetics* 7.3 (1999), pp. 294–301.

[BP89] O. Bíró and K. Preis. "On the use of the magnetic vector potential in the finite-element analysis of three-dimensional eddy currents". In: *IEEE Transactions on Magnetics* 25.4 (1989), pp. 3145–3159.

[Bra+11] H. G. Brachtendorf, T. Tinttunen, W. Schoenmaker, C. Tischendorf, H. H. J. M. Janssen, K. Bittner, E. Dautbegovic, M. Günther, R. Winkler, R. Pulch, A. Schranner, V. Karanko, J. Virtanen, T. Rahkonen, J. Aikio, M. Matthes, S. Baumanns, E. J. W. ter Maten, M. Kole, A. Bouffioux, and D. Hresko. "ICESTARS: Integrated circuit/EM simulation and design technologies for advanced radio systems-on-chip". In: *CASA-report* 1119 (2011).

[Bra+96] H. G. Brachtendorf, G. Welsch, R. Laur, and A. Bunse-Gerstner. "Numerical steady state analysis of electronic circuits driven by multi-tone signals". In: *Electrical Engineering* 79.2 (1996), pp. 103–112.

[Bra01] H. G. Brachtendorf. *Theorie und Analyse von autonomen und quasiperiodisch angeregten elektrischen Netzwerken. Eine algorithmisch orientierte Betrachtung*. Habilitation thesis. Universität Bremen, 2001.

[Bru+17] F. Bruckner, C. Abert, G. Wautischer, C. Huber, C. Vogler, M. Hinze, and D. Suess. "Solving large-scale inverse magnetostatic problems using the adjoint method". In: *Scientific reports* 7 (2017), p. 40816.

[BS17] P. Benner and T. Stykel. "Model order reduction for differential-algebraic equations: A survey". In: *Surveys in Differential-Algebraic Equations IV*. Springer, 2017, pp. 107–160.

[Bur95] K. Burrage. *Parallel and Sequential Methods for Ordinary Differential Equations*. New York, NY, USA: Clarendon Press, 1995.

[But87] J. C. Butcher. *The numerical analysis of ordinary differential equations, Runge–Kutta and general linear methods*. Wiley, Chichester and New York, 1987.

[Cal+00] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari. "Tikhonov regularization and the L-curve for large discrete ill-posed problems". In: *Journal of Computational and Applied Mathematics* 123.1-2 (2000), pp. 423–446.

[Car80] C. J. Carpenter. "Comparison of alternative formulations of 3-dimensional magnetic-field and eddy-current problems at power frequencies". In: *IEE Proceedings B-Electric Power Applications*. Vol. 127. 5. IET. 1980, p. 332.

[Car97] G. F. Carey. *Computational grids: generations, adaptation & solution strategies*. CRC Press, 1997.

[CDK87] L. O. Chua, C. A. Desoer, and E. S. Kuh. *Linear and nonlinear Circuits*. McGraw-Hill Book Co., Singapore, 1987.

[CDS19] I. Cortes Garcia, H. De Gersem, and S. Schöps. "A structural analysis of field/circuit coupled problems based on a generalised circuit element". In: *Numerical Algorithms* 83.1 (Mar. 2019), pp. 373–394.

[CG13] Y. Courvoisier and M. J. Gander. "Optimization of Schwarz waveform relaxation over short time windows". In: *Numerical Algorithms* 64.2 (2013), pp. 221–243.

[CG95] S. L. Campbell and C. W. Gear. "The index of general nonlinear DAEs". In: *Numerische Mathematik* 72 (1995), pp. 173–196.

[CGK75] B. Chawla, H. Gummel, and P. Kozak. “MOTIS-An MOS timing simulator”. In: *IEEE Transactions on Circuits and Systems* 22.12 (1975), pp. 901–910.

[Che+11] Q. Chen, W. Schoenmaker, P. Meuris, and N. Wong. “An effective formulation of coupled electromagnetic-TCAD simulation for extremely high frequency onward”. In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 30.6 (2011), pp. 866–876.

[Che+13] Q. Chen, W. Schoenmaker, G. Chen, L. Jiang, and N. Wong. “A numerically efficient formulation for time-domain electromagnetic-semiconductor cosimulation for fast-transient systems”. In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 32.5 (2013), pp. 802–806.

[Cia02] P. G. Ciarlet. *The finite element method for elliptic problems*. Vol. 40. Siam, 2002.

[CL75] L. O. Chua and P.-M. Lin. “Computer-Aided Analysis of Electronic Circuits: Algorithms and Computational Techniques (Prentice-Hall series in electrical computer engineering)”. In: (1975).

[Cle+02] M. Clemens, H. De Gersem, W. Koch, T. Weiland, and M. Wilke. “Transient simulation of nonlinear electro-quasistatic problems using the finite integration technique”. In: *Proceedings of the 10th International Symposium on Numerical Field Calculations in Electrical Engineering (IGTE 2002)*. 2002, pp. 510–515.

[Cle+11] M. Clemens, S. Schöps, H. De Gersem, and A. Bartel. “Decomposition and regularization of nonlinear anisotropic curl-curl DAEs”. In: *COMPEL-The international journal for computation and mathematics in electrical and electronic engineering* (2011).

[Cle+12] M. Clemens, S. Schöps, C. Cimala, N. Gödel, S. Runke, and D. Schmidthäusler. “Aspects of Coupled Problems in Computational Electromagnetics Formulations”. In: *ICS Newsletter (International Compumag Society)* 19.2 (Nov. 2012), pp. 3–12.

[Cle+20] T. Clees, A. Baldin, P. Benner, S. Grundel, C. Himpe, B. Klaassen, F. Küsters, N. Marheineke, L. Nikitina, I. Nikitin, J. Pade, N. Stahl, C. Strohm, C. Tischendorf, and A. Wirsen. “The MathEnergy Project”. In: *MSO for Power Engineering and Management*. to appear. Springer, 2020.

[Cle05] M. Clemens. “Large systems of equations in a discrete electromagnetism: Formulations and numerical algorithms”. In: *IEE Proceedings-Science, Measurement and Technology* 152.2 (2005), pp. 50–72.

[Cle98] M. Clemens. “Zur numerischen Berechnung zeitlich langsam veränderlicher elektromagnetischer Felder mit der Finiten-Integration-Methode”. PhD thesis. Technische Universität Darmstadt, Fachbereich Elektrotechnik und Informationstechnik, 1998.

[CLP98] J.-F. Charpentier, Y. Lefèvre, and H. Piquet. "An original and natural method of coupling electromagnetic field equations with circuit equations put in a state form". In: *IEEE Transactions on Magnetics* 34.5 (1998), pp. 2489–2492.

[CN08] Z. Chen and J.-C. Nédélec. "On Maxwell Equations with the Transparent Boundary Condition". In: *Journal of Computational Mathematics* 26.3 (2008), pp. 284–296.

[Coo+06] S. J. Cooke, R. Shtokhamer, A. A. Mondelli, and B. Levush. "A finite integration method for conformal, structured-grid, electromagnetic simulation". In: *Journal of Computational Physics* 215.1 (2006), pp. 321–347.

[Cor+17] I. Cortes Garcia, S. Schöps, M. Maciejewski, L. Bortot, M. Prioli, B. Auchmann, and A. Verweij. "Optimized Field/Circuit Coupling for the Simulation of Quenches in Superconducting Magnets". In: *IEEE Journal on Multiscale and Multiphysics Computational Techniques* 2 (2017), pp. 97–104.

[Cor+19] I. Cortes Garcia, S. Schöps, H. De Gersem, and S. Baumanns. "Systems of Differential Algebraic Equations in Computational Electromagnetics". In: *Applications of Differential-Algebraic Equations: Examples and Benchmarks*. Ed. by S. Campbell, A. Ilchmann, V. Mehrmann, and T. Reis. Cham: Springer International Publishing, 2019, pp. 123–169.

[Cor+20] I. Cortes Garcia, S. Schöps, C. Strohm, and C. Tischendorf. "Generalized Elements for a Structural Analysis of Circuits". In: *Progress in Differential-Algebraic Equations II*. Ed. by T. Reis, S. Grundel, and S. Schöps. Cham: Springer International Publishing, 2020, pp. 397–431.

[Cor20] I. Cortes Garcia. "Mathematical Analysis and Simulation of Field Models in Accelerator Circuits". PhD thesis. Technische Universität Darmstadt, Fachbereich Elektrotechnik und Informationstechnik, 2020.

[CW01a] M. Clemens and T. Weiland. "Discrete electromagnetics: Maxwell's equations tailored to numerical simulations". In: *Int. Compumag Soc. Newsletter* 8 (2001), pp. 13–20.

[CW01b] M. Clemens and T. Weiland. "Discrete Electromagnetism with the Finite Integration Technique". In: *Progress In Electromagnetics Research* 32 (2001), pp. 65–87.

[CW02a] M. Clemens and T. Weiland. "Magnetic field simulation using conformal FIT formulations". In: *IEEE Transactions on Magnetics* 38.2 (2002), pp. 389–392.

[CW02b] M. Clemens and T. Weiland. "Regularization of eddy-current formulations using discrete grad-div operators". In: *IEEE Transactions on Magnetics* 38.2 (2002), pp. 569–572.

[De +98] H. De Gersem, R. Mertens, U. Pahner, R. Belmans, and K. Hameyer. "A topological method used for field-circuit coupling". In: *IEEE Transactions on Magnetics* 34.5 (1998), pp. 3190–3193.

[Des81] G. A. Deschamps. "Electromagnetics and differential forms". In: *Proceedings of the IEEE* 69.6 (1981), pp. 676–696.

[DF08] G. Denk and U. Feldmann. "Circuit simulation for nanoelectronics". In: *From Nano to Space*. Springer, 2008, pp. 11–26.

[DGL99] P. Dular, C. Geuzaine, and W. Legros. "A natural method for coupling magnetodynamic H-formulations and circuit equations". In: *IEEE Transactions on Magnetics* 35.3 (1999), pp. 1626–1629.

[DHW04] H. De Gersem, K. Hameyer, and T. Weiland. "Field–circuit coupled models in electromagnetic simulation". In: *Journal of Computational and Applied Mathematics* 168.1-2 (2004), pp. 125–133.

[DK69] C. A. Desoer and E. S. Kuh. *Basic circuit theory*. McGraw-Hill International Editions, 1969.

[DMR05] P. Degond, F. Méhats, and C. Ringhofer. "Quantum energy-transport and drift-diffusion models". In: *Journal of Statistical Physics* 118.3-4 (2005), pp. 625–667.

[DW04] H. De Gersem and T. Weiland. "Field-circuit coupling for time-harmonic models discretized by the finite integration technique". In: *IEEE Transactions on Magnetics* 40.2 (2004), pp. 1334–1337.

[Ebe04] F. Ebert. *Convergence of relaxation methods for coupled systems of ODEs and DAEs*. Tech. rep. urn:nbn:de:0296-matheon-1774. 2004.

[Ebe08] F. Ebert. "On Partitioned Simulation of Electrical Circuits using Dynamic Iteration Methods". PhD thesis. Technische Universität Berlin, 2008.

[Est00] D. Estévez Schwarz. "Consistent initialization for index-2 differential algebraic equations and its application to circuit simulation". PhD thesis. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, 2000.

[Est02] D. Estévez Schwarz. "A step-by-step approach to compute a consistent initialization for the MNA". In: *International Journal of Circuit Theory and Applications* 30.1 (2002), pp. 1–6.

[ET00] D. Estévez Schwarz and C. Tischendorf. "Structural analysis of electric circuits and consequences for MNA". In: *International Journal of Circuit Theory and Applications* 28.2 (2000), pp. 131–162.

[ET88] C. R. I. Emson and C. W. Trowbridge. “Transient 3D eddy currents using modified magnetic vector potentials and magnetic scalar potentials”. In: *IEEE Transactions on Magnetics* 24.1 (1988), pp. 86–89.

[FDW07] M. Funieru, H. De Gersem, and T. Weiland. “Transient Simulation of a Linear Actuator Discretized by the Finite Integration Technique”. In: *Scientific Computing in Electrical Engineering*. Springer, 2007, pp. 281–286.

[FG05] U. Feldmann and M. Günther. “Handbook of Numerical Analysis. Numerical Methods in Electromagnetics”. In: ed. by W. H. A. Schilders and E. J. W. ter Maten. Vol. 13. Elsevier, 2005. Chap. Modelling and Discretization of Circuit Problems, pp. 523–659.

[FG94] U. Feldmann and M. Günther. “The DAE-index in electric circuit simulation”. In: *Proc. IMACS Symposium on Mathematical Modelling*. Ed. by I. Troch and F. Breitenecker. 4. 1994, pp. 695–702.

[Fos92] M. Fosséprez. *Non-linear Circuits: Qualitative Analysis of Non-linear, Non-reciprocal Circuits*. John Wiley & Sons, Chichester, 1992.

[Fri08] A. Friedman. *Partial differential equations of parabolic type*. Courier Dover Publications, 2008.

[Gar05] F. M. Gardner. *Phaselock techniques*. John Wiley & Sons, 2005.

[Gea71] C. W. Gear. “Simultaneous numerical solution of differential-algebraic equations”. In: *IEEE Transactions on Circuit Theory* 18.1 (1971), pp. 89–95.

[GF95] M. Günther and U. Feldmann. “The DAE-index in electric circuit simulation”. In: *Mathematics and Computers in Simulation* 39.5-6 (1995), pp. 573–582.

[GF99] M. Günther and U. Feldmann. “CAD-based electric-circuit modeling in industry I: Mathematical structure and index of network equations”. In: *Surveys on Mathematics for Industry* 8.2 (1999), pp. 97–130.

[Giv91] D. Givoli. “Non-reflecting boundary conditions”. In: *Journal of Computational Physics* 94.1 (1991), pp. 1–29.

[GK96] M. J. Grote and J. B. Keller. “Nonreflecting Boundary Conditions for Time-Dependent Scattering”. In: *Journal of Computational Physics* 127.1 (1996), pp. 52–65.

[GLG85] C. W. Gear, B. Leimkuhler, and G. K. Gupta. “Automatic integration of Euler-Lagrange equations with constraints”. In: *Journal of Computational and Applied Mathematics* 12-13 (May 1985), pp. 77–90.

[GM86] E. Griepentrog and R. März. *Differential–Algebraic Equations and Their Numerical Treatment*. Teubner-Texte zur Mathematik No. 88. BSB B.G. Teubner Verlagsgesellschaft, Leipzig, 1986.

[GP83] C. W. Gear and L. R. Petzold. "Differential-algebraic systems and matrix pencils". In: *Matrix Pencils*. Ed. by B. Kågström and A. Ruhe. Vol. 973. Lecture Notes in Mathematics. Berlin/Heidelberg: Springer, 1983, pp. 75–89.

[GR04] M. J. Gander and A. E. Ruehli. "Optimized waveform relaxation methods for RC type circuits". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 51.4 (2004), pp. 755–768.

[GR10] M. J. Gander and A. E. Ruehli. "Optimized waveform relaxation solution of electromagnetic and circuit problems". In: *19th Topical Meeting on Electrical Performance of Electronic Packaging and Systems*. IEEE. 2010, pp. 65–68.

[GS98] M. J. Gander and A. M. Stuart. "Space-Time Continuous Analysis of Waveform Relaxation for the Heat Equation". In: *SIAM Journal on Scientific Computing* 19.6 (1998), pp. 2014–2031.

[Gün95] M. Günther. "Ladungsorientierte Rosenbrock-Wanner-Methoden zur numerischen Simulation digitaler Schaltungen". PhD thesis. Technische Universität München, 1995.

[Gut98] S. Gutschling. "Zeitbereichsverfahren zur Simulation elektromagnetischer Felder in dispersiven Materialien". PhD thesis. Technische Universität Darmstadt, Jan. 1998.

[GW08] A. Griewank and A. Walther. *Evaluating derivatives: Principles and techniques of algorithmic differentiation*. SIAM, 2008.

[HA01] E. Haber and U. M. Ascher. "Fast finite volume simulation of 3D electromagnetic problems with highly discontinuous coefficients". In: *SIAM Journal on Scientific Computing* 22.6 (2001), pp. 1943–1961.

[Hab04] E. Haber. "Quasi-Newton methods for large-scale electromagnetic inverse problems". In: *Inverse problems* 21.1 (2004), p. 305.

[Hac13] W. Hackbusch. *Multi-grid methods and applications*. Vol. 4. Springer Science & Business Media, 2013.

[Hah92] P. Hahne. "Zur numerischen Feldberechnung zeitharmonischer elektromagnetischer Felder". PhD thesis. Technische Universität Darmstadt, 1992.

[HBG71] G. Hachtel, R. K. Brayton, and F. G. Gustavson. "The Sparse Tableau Approach to Network Analysis and Design". In: *IEEE Transactions on Circuit Theory* 18.1 (1971), pp. 101–113.

[HKO08] R. Hiptmair, F. Kramer, and J. Ostrowski. "A Robust Maxwell Formulation for All Frequencies". In: *IEEE Transactions on Magnetics* 44.6 (2008), pp. 682–685.

[HLR89] E. Hairer, C. Lubich, and M. Roche. *The Numerical Solution of Differential–Algebraic Equations by Runge–Kutta Methods*. Lecture Notes in Mathematics Vol. 1409. Springer-Verlag, Heidelberg, 1989.

[HM76] A. Y. Hannalla and D. C. MacDonald. "Numerical analysis of transient field problems in electrical machines". In: *Proceedings of the Institution of Electrical Engineers*. Vol. 123. 9. IET. 1976, pp. 893–898.

[HM89] H. A. Haus and J. R. Melcher. *Electromagnetic fields and energy*. Vol. 107. Prentice Hall Englewood Cliffs, NJ, 1989.

[HN85] E. E. E. Hoefer and H. Nielinger. *SPICE*. Springer Berlin Heidelberg, 1985.

[HRB75] C.-W. Ho, A. E. Ruehli, and P. Brennan. "The modified nodal approach to network analysis". In: *IEEE Trans. Circuits Syst.* 22.6 (June 1975), pp. 504–509.

[Huc18] C. Huck. "Perturbation analysis and numerical discretisation of hyperbolic partial differential algebraic equations describing flow networks". PhD thesis. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2018.

[HW91] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential–Algebraic Problems*. Springer Series in Computational Mathematics 14. Springer-Verlag, Berlin, Heidelberg, 1991.

[Jac99] J. D. Jackson. *Classical Electrodynamics*. 3rd ed. John Wiley, 1999.

[Jan15] L. Jansen. "A dissection concept for DAEs: Structural decoupling, unique solvability, convergence theory and half-explicit methods." PhD thesis. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät I, 2015.

[Jer12] J. W. Jerome. *Analysis of charge transport: A mathematical study of semiconductor devices*. Springer Science & Business Media, 2012.

[JK96] Z. Jackiewicz and M. Kwapisz. "Convergence of Waveform Relaxation Methods for Differential-Algebraic Systems". In: *SIAM Journal on Numerical Analysis* 33.6 (Dec. 1996), pp. 2303–2317.

[JMT15] L. Jansen, M. Matthes, and C. Tischendorf. "Global unique solvability for memristive circuit DAEs of index 1". In: *International Journal of Circuit Theory and Applications* 43.1 (2015), pp. 73–93.

[JS95] K. Jbilou and H. Sadok. "Analysis of some vector extrapolationmethods for solving systems of linear equations". In: *Numerische Mathematik* 70.1 (Mar. 1995), pp. 73–89.

[Kam90] A. Kameari. “Calculation of transient 3D eddy current using edge-elements”. In: *IEEE Transactions on Magnetics* 26.2 (1990), pp. 466–469.

[Kan01] S. Kanerva. “Data transfer methodology between a FEM program and a system simulator”. In: *ICEMS’2001. Proceedings of the Fifth International Conference on Electrical Machines and Systems (IEEE Cat. No. 01EX501)*. Vol. 2. IEEE. 2001, pp. 1121–1124.

[KM06] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations. Analysis and Numerical Solutions*. European Mathematical Publishing House, Feb. 2006.

[KM94] P. Kunkel and V. Mehrmann. “Canonical forms for linear differential–algebraic equations with variable coefficients”. In: *Journal of Computational and Applied Mathematics* 56.3 (1994), pp. 225–251.

[Krü00] H. Krüger. *Zur numerischen Berechnung transienter elektromagnetischer Felder in gyrotropen Materialien*. Der Andere Verlag, 2000.

[LBH84] R. W. Lewis, P. Bettess, and E. Hinton, eds. *Numerical methods in coupled systems*. Numerical Methods in Engineering. John Wiley, Chichester, 1984.

[Lel82] E. Lelarasmee. *The waveform relaxation method for time domain analysis of large scale integrated circuits: Theory and applications*. Electronics Research Laboratory, College of Engineering, University of California, 1982.

[LLC97] J.-F. Lee, R. Lee, and A. Cangellaris. “Time-domain finite-element methods”. In: *IEEE Transactions on Antennas and Propagation* 45.3 (1997), pp. 430–442.

[LM93] P. Lombard and G. Meunier. “A general purpose method for electric and magnetic combined problems for 2D, axisymmetric and transient systems”. In: *IEEE Transactions on Magnetics* 29.2 (1993), pp. 1737–1740.

[LMT05] R. Lamour, R. März, and C. Tischendorf. *PDAEs and further mixed systems as abstract differential algebraic systems*. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2005.

[LMT13] R. Lamour, R. März, and C. Tischendorf. *Differential-Algebraic Equations: A Projector Based Analysis*. Springer, Heidelberg, 2013.

[LP16] H. P. Langtangen and G. K. Pedersen. *Scaling of differential equations*. Springer Nature, 2016.

[LRS82] E. Lelarasmee, A. E. Ruehli, and A. L. Sangiovanni-Vincentelli. “The waveform relaxation method for time-domain analysis of large scale integrated circuits”. In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 1.3 (July 1982), pp. 131–145.

[Lun00] M. Lundstrom. *Fundamentals of Carrier Transport*. Cambridge University Press (CUP), 2000.

[MA87] G. Massabrio and P. Antognetti. *Semiconductor Device Modeling with Spice*. Ed. by G. Massabrio and P. Antognetti. McGraw-Hill Book Co., Singapore, 1987.

[MAG16] MAGWEL N.V. *Device-Electro-Magnetic Modeler (DevEM^{TM})*. 2016.

[Mar02] R. Marklein. *The finite integration technique as a general tool to compute acoustic, electromagnetic, elastodynamic, and coupled wave fields*. IEEE Press, John Wiley, and Sons, New York, USA, 2002.

[Mär03] R. März. "Differential–algebraic systems with properly stated leading term and MNA equations". In: *Modelling, Simulation and Optimization of Integrated Circuits*. Ed. by K. Anstreich, R. Bulirsch, A. Gilg, and P. Rentrop. Birkhäuser, 2003, pp. 135–151.

[Mar86] P. A. Markowich. *The Stationary Semiconductor Device Equations*. Springer Vienna, 1986.

[Mär87] R. März. *A matrix chain for analyzing differential–algebraic equations*. Preprint (Neue Folge) 162. Humboldt-Universtät zu Berlin, Sektion Mathematik, 1987.

[Mat+16] E. J. W. ter Maten, P. A. Putek, M. Günther, R. Pulch, C. Tischendorf, C. Strohm, W. Schoenmaker, P. Meuris, B. De Smedt, P. Benner, et al. "Nanoelectronic COupled problems solutions-nanoCOPS: modelling, multirate, model order reduction, uncertainty quantification, fast fault simulation". In: *Journal of Mathematics in Industry* 7.1 (2016), p. 2.

[Mat+19] E. J. W. ter Maten, H. G. Brachtendorf, R. Pulch, W. Schoenmaker, and H. De Gersem. *Nanoelectronic Coupled Problems Solutions*. Vol. 29. Mathematics in Industry. Springer Nature, 2019.

[Mat12] M. Matthes. *Numerical Analysis of Nonlinear Partial Differential-Algebraic Equations. A Coupled and an Abstract Systems Approach*. Logos Verlag, Berlin, 2012.

[Mat96] I. MathWorks. *SIMULINK dynamic system simulation for MATLAB: Modeling, simulation, implementation*. Vol. 1. MathWorks, Inc., 1996.

[Max65] J. C. Maxwell. "A Dynamical Theory of the Electromagnetic Field". In: *Philosophical Transactions of the Royal Society* 155 (1865), pp. 459–512.

[Mer+98] R. Mertens, H. De Gersem, R. Belmans, K. Hameyer, D. Lahaye, S. Vandewalle, and D. Roose. "An algebraic multigrid method for solving very large electromagnetic systems". In: *IEEE Transactions on Magnetics* 34.5 (1998), pp. 3327–3330.

[Mie89] U. Miekkala. “Dynamic iteration methods applied to linear DAE systems”. In: *Journal of Computational and Applied Mathematics* 25.2 (1989), pp. 133–151.

[ML09] T. G. Mackay and A. Lakhtakia. *Electromagnetic Anisotropy and Bianisotropy*. World Scientific Pub Co Pte Lt, Nov. 2009.

[MN87] U. Miekkala and O. Nevanlinna. “Convergence of Dynamic Iteration Methods for Initial Value Problems”. In: *SIAM Journal on Scientific and Statistical Computing* 8.4 (1987), pp. 459–482.

[MNS17] M. Merkel, I. Niyonzima, and S. Schöps. “ParaExp Using Leapfrog as Integrator for High-Frequency Electromagnetic Simulations”. In: *Radio Science* 52.12 (2017), pp. 1558–1569.

[MSM01] P. Meuris, W. Schoenmaker, and W. Magnus. “Strategy for electromagnetic interconnect modeling”. In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 20.6 (June 2001), pp. 753–762.

[MW07] I. Munteanu and T. Weiland. “RF & microwave simulation with the finite integration technique–from component to system design”. In: *Scientific Computing in Electrical Engineering*. Springer, 2007, pp. 247–260.

[Nag75] L. W. Nagel. “SPICE2: A computer program to simulate semiconductor circuits”. PhD thesis. University of California at Berkeley, 1975.

[Nic14] S. Nicaise. “Existence results for the A - phi magnetodynamic formulation of the Maxwell system”. In: *Applicable Analysis* 94.5 (2014), pp. 1–16.

[Nol11] W. Nolting. *Grundkurs Theoretische Physik 3*. Springer Berlin Heidelberg, 2011.

[NP73] L. W. Nagel and D. O. Pederson. *SPICE (Simulation Program with Integrated Circuit Emphasis)*. Tech. rep. UCB/ERL M382. EECS Department, University of California, Berkeley, Apr. 1973.

[OOL04] J. T. Ottesen, M. S. Olufsen, and J. K. Larsen. *Applied Mathematical Models in Human Physiology*. Society for Industrial & Applied Mathematics (SIAM), Jan. 2004.

[Pad20] J. Pade. “Analysis and waveform relaxation for a differential-algebraic electrical circuit model”. submitted. PhD thesis. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät I, 2020.

[PCW03] O. Podebrad, M. Clemens, and T. Weiland. “New flexible subgridding scheme for the finite integration technique”. In: *IEEE Transactions on Magnetics* 39.3 (2003), pp. 1662–1665.

[PG02] R. Pulch and M. Günther. "A method of characteristics for solving multirate partial differential equations in radio frequency application". In: *Applied numerical mathematics* 42.1-3 (2002), pp. 397–409.

[Pio93] A. Piotrowska. "Ohmic Contacts to GaAs: Fundamentals and Practice". In: *International School of Semiconducting Compounds*. Vol. 84. 3. Institute of Physics, Polish Academy of Sciences, Sept. 1993, pp. 491–504.

[PT18] J. Pade and C. Tischendorf. "Waveform relaxation: A convergence criterion for differential-algebraic equations". In: *Numerical Algorithms* (Dec. 2018).

[Pus+14] B. Pussig, J. Denil, P. De Meulenaere, and H. Vangheluwe. "Generation of functional mock-up units for co-simulation from simulink®, using explicit computational semantics: Work in progress paper". In: *Proceedings of the Symposium on Theory of Modeling & Simulation-DEVS Integrative*. 2014, pp. 1–6.

[QSS10] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical mathematics*. Vol. 37. Springer Science & Business Media, 2010.

[QV08] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*. Vol. 23. Springer Science & Business Media, 2008.

[Rei98] M. Reisch. *Elektronische Bauelemente*. Springer Science + Business Media, 1998.

[Rie01] U. van Rienen. *Numerical Methods in Computational Electrodynamics: Linear Systems in Practical Applications*. Springer Science & Business Media, 2001.

[RR02] P. J. Rabier and W. C. Rheinboldt. "Techniques of scientific computing (Part 4) - Theoretical and Numerical Analysis of Differential–Algebraic Equations". In: *Handbook of Numerical Analysis*. Ed. by P. G. Ciarlet and K. L. Lions. Vol. VIII. Amsterdam: North Holland/ Elsevier, 2002, pp. 183–540.

[RS09] J. Rommes and W. H. A. Schilders. "Efficient methods for large resistor networks". In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 29.1 (2009), pp. 28–39.

[RSM90] C. A. Ringhofer, C. Schmeiser, and P. A. Markowich. *Semiconductor Equations*. 1990.

[Saa03] Y. Saad. *Iterative methods for sparse linear systems*. SIAM, 2003.

[Sau19] H. Sauter. "Waveform-Relaxations-Verfahren zur numerischen Simulation gekoppelter Differential-algebraischer Gleichungen". MA thesis. Humboldt Universität zu Berlin, 2019.

[SBD12] S. Schöps, A. Bartel, and H. De Gersem. “Multirate Time Integration of Field/Circuit Coupled Problems by Schur Complements”. In: *Scientific Computing in Electrical Engineering SCEE 2010*. Springer, 2012, pp. 243–251.

[SCG14] W. Schoenmaker, Q. Chen, and P. Galy. “Computation of self-induced magnetic field effects including the lorentz force for fast-transient phenomena in integrated-circuit devices”. In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 33.6 (2014), pp. 893–902.

[Sch+10a] S. Schöps, A. Bartel, H. De Gersem, and M. Günther. “DAE-Index and Convergence Analysis of Lumped Electric Circuits Refined by 3-D Magnetoquasistatic Conductor Models”. In: *Scientific Computing in Electrical Engineering SCEE 2008*. Ed. by J. Roos and L. R. Costa. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 341–348.

[Sch+10b] S. Schöps, A. Bartel, H. De Gersem, and M. Günther. “DAE-index and convergence analysis of lumped electric circuits refined by 3-d magnetoquasistatic conductor models”. In: *Scientific Computing in Electrical Engineering SCEE 2008*. Springer, 2010, pp. 341–348.

[Sch+16] W. Schoenmaker, P. Meuris, C. Strohm, and C. Tischendorf. “Holistic Coupled Field and Circuit Simulation”. In: *Proceedings of the 2016 Conference on Design, Automation & Test in Europe*. DATE ’16. EDA Consortium. Dresden, Germany: EDA Consortium, 2016, pp. 307–312.

[Sch+19a] W. Schoenmaker, H. G. Brachtendorf, K. Bittner, C. Tischendorf, and C. Strohm. “Discretizations”. In: *Nanoelectronic Coupled Problems Solutions*. Springer, 2019, pp. 69–91.

[Sch+19b] W. Schoenmaker, H. G. Brachtendorf, K. Bittner, C. Tischendorf, and C. Strohm. “Holistic/Monolithic Time Integration”. In: *Nanoelectronic Coupled Problems Solutions*. Springer, 2019, pp. 117–130.

[Sch11] S. Schöps. “Multiscale modeling and multirate time-integration of field/circuit coupled problems”. PhD thesis. Universität Wuppertal, Fakultät für Mathematik und Naturwissenschaften Mathematik und Informatik Dissertationen, 2011.

[Sch99] R. Schuhmann. “Die Nichtorthogonale Finite-Integrations-Methode zur Simulation elektromagnetischer Felder”. PhD thesis. Technische Universität Darmstadt, 1999.

[SDB10] S. Schöps, H. De Gersem, and A. Bartel. “A Cosimulation Framework for Multirate Time Integration of Field/Circuit Coupled Problems”. In: *IEEE Transactions on Magnetics* 46.8 (Aug. 2010), pp. 3233–3236.

[SDB12] S. Schöps, H. De Gersem, and A. Bartel. “Higher-order cosimulation of field/circuit coupled problems”. In: *IEEE Transactions on Magnetics* 48.2 (2012), pp. 535–538.

[SDW13] S. Schöps, H. De Gersem, and T. Weiland. “Winding functions in transient magnetoquasistatic field-circuit coupled simulations”. In: *COMPEL: The international journal for computation and mathematics in electrical and electronic engineering* (2013).

[Sel84] S. Selberherr. *Analysis and Simulation of Semiconductor Devices*. Springer Nature, 1984.

[Ske80] R. D. Skeel. “Iterative refinement implies numerical stability for Gaussian elimination”. In: *Mathematics of Computation* 35.151 (1980), pp. 817–832.

[SM05] W. H. A. Schilders and E. J. W. ter Maten. *Handbook of Numerical Analysis: Special Volume Numerical Methods in Electrodynamics*. Ed. by P. Ciarlet. Vol. XIII. Elsevier North Holland, 2005.

[SSH08] K. Schmidt, O. Sterz, and R. Hiptmair. “Estimating the eddy-current modeling error”. In: *IEEE Transactions on Magnetics* 44.6 (2008), pp. 686–689.

[ST20] C. Strohm and C. Tischendorf. “Coupled Electromagnetic Field and Electric Circuit Simulation: A Waveform Relaxation Benchmark”. In: *Modeling, Simulation and Optimization of Complex Processes HPSC 2018: Proceedings of the Seventh International Conference on High Performance Scientific Computing, March 19-23, 2018, Hanoi, Vietnam*. Ed. by H. G. Bock, H. X. Phu, R. Rannacher, and J. P. Schlöder. Springer, 2020.

[Ste+08] T. Steinmetz, N. Godel, G. Wimmer, M. Clemens, S. Kurz, and M. Bebendorf. “Efficient symmetric FEM-BEM coupled simulations of electro-quasistatic fields”. In: *IEEE Transactions on Magnetics* 44.6 (2008), pp. 1346–1349.

[Str+18] T. Streubel, C. Strohm, P. Trunschke, and C. Tischendorf. “Generic Construction and Efficient Evaluation of Flow Network DAEs and Their Derivatives in the Context of Gas Networks”. In: *Operations Research Proceedings 2017*. Springer, 2018, pp. 627–632.

[Str07] J. A. Stratton. *Electromagnetic Theory*. Wiley-Blackwell, Oct. 2007.

[SVR08] W. H. A. Schilders, H. A. Van der Vorst, and J. Rommes. *Model order reduction: Theory, research aspects and applications*. Vol. 13. Springer, 2008.

[SW00] R. Schuhmann and T. Weiland. “The nonorthogonal finite integration technique applied to 2D-and 3D-eigenvalue problems”. In: *IEEE Transactions on Magnetics* 36.4 (2000), pp. 897–901.

[SW01] R. Schuhmann and T. Weiland. “Conservation of Discrete Energy and Related Laws in the Finite Integration Technique”. In: *Progress in Electromagnetics Research* 32 (2001), pp. 301–316.

[SW98a] R. Schuhmann and T. Weiland. “A stable interpolation technique for FDTD on non-orthogonal grids”. In: *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields* 11.6 (1998), pp. 299–306.

[SW98b] R. Schuhmann and T. Weiland. “Stability of the FDTD algorithm on nonorthogonal grids related to the spatial interpolation scheme”. In: *IEEE Transactions on Magnetics* 34.5 (1998), pp. 2751–2754.

[SW99] R. Schuhmann and T. Weiland. “FDTD on nonorthogonal grids with triangular fillings”. In: *IEEE Transactions on Magnetics* 35.3 (1999), pp. 1470–1473.

[SZF06] W. Sun, J.-H. Zou, and X.-G. Fan. “Convergence of parallel dynamic iteration methods for nonlinear DAEs of index-2”. In: *2006 IEEE International Conference on Automation Science and Engineering*. IEEE. 2006, pp. 129–133.

[Tho97] P. Thoma. “Zur numerischen Lösung der Maxwellschen Gleichungen im Zeitbereich”. PhD thesis. Technische Universität Darmstadt, Fachbereich Elektrotechnik und Informationstechnik, 1997.

[Tis04] C. Tischendorf. *Coupled Systems of Differential Algebraic and Partial Differential Equations in Circuit and Device Simulation*. Habilitation thesis. Humboldt Universität zu Berlin, 2004.

[Tis99] C. Tischendorf. “Topological index calculation of DAEs in circuit simulation”. In: *Surv. Math. Ind.* 8.3-4 (1999), pp. 187–199.

[Ton01] E. Tonti. “A direct discrete formulation of field laws: The cell method”. In: *CMES-Computer Modeling in Engineering and Sciences* 2.2 (2001), pp. 237–258.

[Ton75] E. Tonti. *On the formal structure of physical theories*. Istituto di matematica del Politecnico di Milano, 1975.

[Ton95] E. Tonti. “On the geometrical structure of electromagnetism”. In: *Gravitation, Electromagnetism and Geometrical Structures* (1995), pp. 281–308.

[Tsu+93] I. A. Tsukerman, A. Konrad, G. Meunier, and J. C. Sabonnadiere. “Coupled field-circuit problems: Trends and accomplishments”. In: *IEEE Transactions on Magnetics* 29.2 (1993), pp. 1701–1704.

[Tsu02] I. A. Tsukerman. “Finite element differential–algebraic systems for eddy current problems”. In: *Numerical Algorithms* 31.1-4 (2002), pp. 319–335.

[TW96] P. Thoma and T. Weiland. “A consistent subgridding scheme for the finite difference time domain method”. In: *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields* 9.5 (1996), pp. 359–374.

[UKM95] J. Unger, A. Kröner, and W. Marquardt. "Structural analysis of Differential–Algebraic Equation Systems - Theory and Applications". In: *Computers chem. Engng* 19.8 (1995), pp. 867–882.

[Vaa96] J. Vaananen. "Circuit theoretical approach to couple two-dimensional finite element models with external circuit equations". In: *IEEE Transactions on Magnetics* 32.2 (1996), pp. 400–410.

[Vol09] V. I. Voloshin. *Introduction to Graph and Hypergraph Theory*. NOVA Science Publishers, 2009.

[Wan96] J.-S. Wang. "A nodal analysis approach for 2D and 3D magnetic-circuit coupled problems". In: *IEEE Transactions on Magnetics* 32.3 (1996), pp. 1074–1077.

[Wei77a] T. Weiland. "A Discretization Method for the Solution of Maxwell's Equations for Six-Component Fields". In: *AEÜ - International Journal of Electronics and Communications* 31 (1977), pp. 166–120.

[Wei77b] T. Weiland. "Zur numerischen Lösung des Eigenwellenproblems längshomogener Wellenleiter beliebiger Randkontur und transversal inhomogener Füllung". PhD thesis. Technische Universität Darmstadt, Fachbereich Elektrotechnik und Informationstechnik, 1977.

[Wei79] T. Weiland. "Lossy waveguides with arbitrary boundary contour and material distribution". In: *Archiv Elektronik und Uebertragungstechnik* 33 (1979), pp. 170–174.

[Wei84] T. Weiland. "On the numerical solution of Maxwell's equations and applications in the field of accelerator physics". In: *Part. Accel.* 15.DESY-84-006 (1984), pp. 245–292.

[Wei96] T. Weiland. "Time domain electromagnetic field computation with finite difference methods". In: *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields* 9.4 (1996), pp. 295–319.

[WH06] A. Wachter and H. Hoeber. *Compendium of Theoretical Physics*. Springer Science & Business Media, 2006.

[Whi+85] J. K. White, F. Odeh, A. L. Sangiovanni-Vincentelli, and A. E. Ruehli. *Waveform Relaxation: Theory and Practice*. Memorandum UCB/ERL M85/65. 1985.

[WL03] W. S. Wieglhofer and A. Lakhtakia, eds. *Introduction to Complex Mediums for Optics and Electromagnetics*. SPIE-Intl Soc Optical Eng, Mar. 2003.

[WS12] J. K. White and A. L. Sangiovanni-Vincentelli. *Relaxation techniques for the simulation of VLSI circuits*. Vol. 20. Springer Science & Business Media, 2012.

[Yee66] K. Yee. “Numerical solution of initial boundary value problems involving Maxwell’s equations in isotropic media”. In: *IEEE Transactions on antennas and propagation* 14.3 (1966), pp. 302–307.

[Zei13] E. Zeidler. *Nonlinear Functional Analysis and Its Applications: II/B: Nonlinear Monotone Operators*. Springer Science & Business Media, 2013.

[Zei90] E. Zeidler. *Nonlinear Functional Analysis and its Applications II/B: Nonlinear Monotone Operators*. Springer, 1990.

[Zho+06] P. Zhou, D. Lin, W. N. Fu, B. Ionescu, and Z. J. Cendes. “A general cosimulation approach for coupled field-circuit problems”. In: *IEEE Transactions on Magnetics* 42.4 (2006), pp. 1051–1054.

[Zho06] X. Zhou. “On independence, completeness of Maxwell’s equations and uniqueness theorems in electromagnetics”. In: *Progress in Electromagnetics Research* 64 (2006), pp. 117–134.

Index

backward differentiation formula, 36, 113
balun circuit, 157
Banach fixed point theorem, 139
band-pass filter, 158, 167
Boltzmann constant, 22, 42
boundary condition
    artificial, 46–48
    discrete, 74, 77
    electric, 47, 51
    magnetic, 47
    physical, 46
boundary conditions, 43
    artificial, 43
    Dirichlet, 47, 51
    Neumann, 47
    physical, 43
branch, 16
    current, 20, 85
    voltage, 20, 85

cell complex, 55
    dual, 58
    orthogonal staggered, 60, 63
cell method, 36, 54
characteristic function, 25
charge conservation, 39
co-simulation, 109
conductivity, 41
    matrix, 72
conjugate gradient, 123
consistency, 25
consistent initial value, 6, 25
constitutive element equations, 21
constitutive grid relations, 73
constitutive relations, 35, 40, 41
continuity equation, 40
convection current density, 39
convergence
    Gauss-Seidel type, 141
    Jacobi type, 150
    waveform relaxation, 135
coupled system, 9, 153
    additive coupled, 10
    black-box versions, 95, 130, 133
    field/circuit analysis, 97
    field/circuit modeling, 82
    incorporated versions, 94, 127, 132
    index, 101, 105
    Jacobian, 116
    shifted versions, 94, 129, 133
    waveform relaxation, 125
curl matrix, 67
    reduced, 77
current continuity equation, 38, 41
current controlling, 21
cutset, 17, 182
    I-, 25
    IM-, 26
    LIE^{+}-, 147, 150, 164

differential-algebaic equation
    matrix chain, 98
differential-algebraic equation
    in standard form, 6
    matrix chain, 12
    properly stated derivative term, 7
    semi-properly stated derivative term, 8
    with nonlinear derivative term, 6
discrete operators, 67
    reduced, 77
distribution of charges, 66
divergence matrix, 67

reduced, 77
doping profile, 41
drift-diffusion model, 35, 41

eddy-current problems, 35
electric charge density, 37
electric circuit, 16
modeling, 20
proper, 17
electric current density, 37, 39, 41
electric field strength, 37
electric flux density, 37
electric mesh currents, 66
electric mesh fluxes, 66
electric mesh voltages, 66
electric scalar potential, 49
discrete, 75
electromagnetic device, 35
modeling, 41, 52, 78
electron concentration, 41
element
mock, 15, 87
multi-terminal, 16
two-terminal, 21
elementary charge, 22, 41
equivalent circuit, 23
external source densities, 39

finite difference time domain, 36, 54
finite element method, 36, 54
finite integration technique, 36, 37, 54
Functional Mock-up Interface, 15

gauge condition, 51
Coulomb, 51
grad-type Lorenz, 51
Lorenz, 51
Gauss's law, 37
discrete, 75
Gauss's law for magnetism, 37
generalized minimal residual method, 117
geometrical objects, 55
gradient matrix, 67
reduced, 77
graph, 16, 182
connected, 17
hyper-, 16

hidden constraints, 5
hole concentration, 41

ideality factor, 22
idealized surface charge density, 44
idealized surface density, 45
implicit Euler method, 8, 114
incidence matrix, 17
full, 17
reduced, 17, 84
index
differentiation index, 5
dissection index, 5, 10, 97, 101
perturbation index, 5
strangeness index, 5
tractability index, 5
index reduction
Gear-Gupta-Leimkuhler, 6
initial value problem, 6
interface conditions, 43

Jacobian, 115
jet-derivative operator, 12

kernel splitting pair, 11, 99
Kirchhoff's circuit laws, 20, 40
current law, 20
voltage law, 20

leapfrog, 36
linear solving
direct, 117
hybrid, 117
iterative, 117
Lipschitz continuous, 135
loop, 17, 181
V-, 25, 149
XY^+-, 18
low-pass filter, 153, 163

magnetic field strength, 37
magnetic flux density, 37
magnetic mesh fluxes, 66
magnetic mesh voltages, 66

magnetic vector potential, 49
    discrete, 75
material, 41, 51
    matrices, 72
    properties, 153
Maxwell's equations, 35, 37
    $A - \varphi$ formulation, 49, 50
    differential form, 38
    integral form, 38
Maxwell's grid equations, 36, 54, 65
    $A - \varphi$ formulation, 75
Maxwell's house, 50
Maxwell-Ampère's law, 37
    discrete, 75
Maxwell-Faraday's law, 37
media
    conducting, 41, 52
    insulating, 41, 52
    semi-conducting, 41
mesh
    computational, 55
    rectilinear, 56
modified nodal analysis, 24
model order reduction, 15, 112
monolithic, 109
    benchmark, 153

Nemytskii operator, 136
Newton's method, 113, 115
node, 16
    ground, 17, 84
    non-reference, 16
    potential, 20
    reference, 16
nominal temperature, 22, 42

Ohm's law, 40

p-cell, 55
    orientation, 61
passivity, 25
permeability, 41
    of vacuum, 51
permittivity, 41
    matrix, 72
    of vacuum, 51
principle of causality, 40

quasi-canonical momentum, 50, 122
    discrete, 78, 123
quasistatic
    electro-, 35
    magneto-, 35

reluctivity, 41
    matrix, 72
reverse biased saturation current, 22
Runge-Kutta methods, 36, 114

scaling, 119
    equation, 121
    row, 121
    time, 121
    variable, 120
Schur complement, 117, 122
semiconductor, 160
    modeling, 41, 52
semiconductor modeling, 42
Shockley equation, 22
smoothness, 25

test cases, 157
thermal voltage, 22
Tonti's diagram, 50
total current
    density, 39, 85
    flow, 76, 85
tree, 17, 182

voltage controlling, 21

waveform relaxation, 110, 124, 135
    benchmark, 162
    Gauss-Seidel type, 125, 126
    Gauss-Seidel type benchmark, 163
    Jacobi type, 126, 132
    Jacobi type benchmark, 163
window functions, 62
windowing technique, 126, 169